26925

National Library of Canada | Bibliothèque nationale du Canada

CANADIAN THESES ON MICROFICHE | THÈSES CANADIENNES SUR MICROFICHE

NAME OF AUTHOR/*NOM DE L'AUTEUR* JAMES DALE STAFFORD

TITLE OF THESIS/*TITRE DE LA THÈSE* DIFFERENTIAL URBAN DEVELOPMENT IN CANADA, 1951-1971

UNIVERSITY/*UNIVERSITÉ* UNIVERSITY OF ALBERTA

DEGREE FOR WHICH THESIS WAS PRESENTED/ *GRADE POUR LEQUEL CETTE THÈSE FUT PRÉSENTÉE* Ph. D.

YEAR THIS DEGREE CONFERRED/*ANNÉE D'OBTENTION DE CE GRADE* 1975

NAME OF SUPERVISOR/*NOM DU DIRECTEUR DE THÈSE* Dr. Karol J. Krotki

Permission is hereby granted to the NATIONAL LIBRARY OF CANADA to microfilm this thesis and to lend or sell copies of the film.

The author reserves other publication rights, and neither the thesis nor extensive extracts from it may be printed or otherwise reproduced without the author's written permission.

L'autorisation est, par la présente, accordée à la BIBLIOTHÈQUE NATIONALE DU CANADA de microfilmer cette thèse et de prêter ou de vendre des exemplaires du film.

L'auteur se réserve les autres droits de publication; ni la thèse ni de longs extraits de celle-ci ne doivent être imprimés ou autrement reproduits sans l'autorisation écrite de l'auteur.

DATED/*DATÉ* 17 Oct. 1975 SIGNED/*SIGNÉ* J. Stafford

PERMANENT ADDRESS/*RÉSIDENCE FIXE*
Dept. of Sociology
Lakehead University
Thunder Bay, Ontario P7B 5E1

THE UNIVERSITY OF ALBERTA

RELEASE FORM

NAME OF AUTHOR James Dale Stafford

TITLE OF THESIS Differential Urban Development in Canada, 1951-1961

DEGREE FOR WHICH THESIS WAS PRESENTED Ph.D.

YEAR THIS DEGREE GRANTED 1975

Permission is hereby granted to THE UNIVERSITY OF ALBERTA LIBRARY to reproduce single copies of this thesis and to lend or sell such copies for private, scholarly or scientific research purposes only.

The author reserves other publication rights, and neither the thesis nor extensive extracts from it may be printed or otherwise reproduced without the author's written permission.

J. Stafford

PERMANENT ADDRESS:

Lakehead University
Thunder Bay, Ontario
P7B 5E1

DATED October 4, 1975

THE UNIVERSITY OF ALBERTA

DIFFERENTIAL URBAN DEVELOPMENT IN CANADA, 1951-1961

by

JAMES DALE STAFFORD

A THESIS
SUBMITTED TO THE FACULTY OF GRADUATE STUDIES AND RESEARCH
IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

DEPARTMENT OF SOCIOLOGY

EDMONTON, ALBERTA
FALL, 1975

THE UNIVERSITY OF ALBERTA

FACULTY OF GRADUATE STUDIES AND RESEARCH

The undersigned certify that they have read, and recommend to the Faculty of Graduate Studies and Research, for acceptance, a thesis entitled "Differential Urban Growth in Canada, 1951 - 1961" submitted by James Dale Stafford in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Sociology.

....
Supervisor

W. Gordon Hamel

....

R.G. Ironside

Joel W. Gregory
....
External Examiner

Date 25 September 1975

ABSTRACT

The major concern is the identification of factors leading to the differential urban growth of Canadian cities from 1951 to 1961. This thesis attempts to answer the question: why do certain cities develop more rapidly than others? No adequate theory of differential urban growth exists, so three theories were borrowed from regional science and compared with respect to changes in the Canadian urban system from 1951 to 1961.

The three theories of regional growth are sector theory, export base theory, and growth pole theory. Sector theory states that a region develops as the major components of its economy shift from primary to secondary and from secondary to tertiary industries. According to export base theory, a region develops because it contains a commodity which is in demand outside the region. Growth pole theory attributes development to agglomeration, industrial linkages, and entrepreneurial activity.

Variables that are considered to be essential to each theory are operationalized. The corresponding data for the forty-six largest cities in Canada are regressed on migration

and wage data using a stepwise multiple regression model. The results indicate that the variables representing export base theory contribute very little to reduction of the unexplained variance in the dependent variables. Sector theory contributes more, and growth pole theory most, to reduction of this variance. Two variables from growth pole theory, agglomeration and profits, are most powerful in accounting for variation in wages and migration. The most powerful variable from sector theory is percentage of the manufacturing labour force engaged in capital-goods industries.

A theory of differential urban development in Canada must extend beyond traditional economic variables. Other factors, as indicated by studies of Canada's urban system include regional characteristics. Associated with regionalism is the heartland-hinterland concept and the economic shadow concept, both of which favor cities in southern Ontario. The financial central place functions of some large cities, particularly Toronto and Montreal, must also be considered. Finally the advantages of Metropolitan Areas and capital cities play a role in differential urban development.

ACKNOWLEDGEMENTS

I wish to offer special thanks to Dr. Krotki who spent so many hours reading and criticizing the various drafts of my proposal and thesis. Without him, this thesis would have contained many more errors than it now does. His standards stimulated me to perform at a level higher than my normal lethargy allows.

I also wish to extend my gratitude to Dr. Krishnan and Dr. Harrel for their advice with respect to the methodology of the thesis. Thanks are extended to Dr. Proudfoot and Dr. Ironside for gently guiding me through the alien but interesting discipline of economic geography. I appreciate the patience they exhibited and the advice they gave while I grappled with some of the basic concepts in that field.

An expression of thanks is due Susan Racine and Zoe Bonnett for the many hours they spent typing the final manuscript. They also deserve an apology for the many frustrations I caused them.

And Julie.

TABLE OF CONTENTS

CHAPTER PAGE

1. INTRODUCTION 1
    - Urbanization 2
    - Urban Growth 5
    - Causes of Urban Growth 7
    - Outline of the Remainder of the Thesis . . 20

2. THEORIES 22
    - Export Base Theory 22
    - Research 32
    - Criticisms 34
    - Sector Theory 38
    - Research 43
    - Criticisms 45
    - Growth Pole Theory 49
    - Research 57
    - Criticisms 63
    - Comparisons 66
    - Summary 73

3. METHODOLOGY 77
    - Operational Definitions of the Export Base 78
    - Operational Definitions of Industrial Sectors 93
    - Operational Definitions of Crucial Variables in Growth Pole Theory 95
    - Unit of Analysis 98
    - Operational Definition of Urban Development 104
    - Operational Definition of Change 107
    - The Model 113
    - Summary 114

4. DATA ANALYSIS 118
    - Migration Impact 156
    - Average Value Added 161
    - Further Analysis 164
    - Summary 175

TABLE OF CONTENTS

CHAPTER PAGE

5. DISCUSSION AND CONCLUSIONS 180
   Discussion of Major Variables 180
   Contributions to a Theory of Differential Urban Development in Canada 207
   Regionalism 208
   Capital Cities 234
   Metropolitan Areas 235
   Heartland-Hinterland 236
   Economic Shadow 242
   Financial Structure 243
   Problems for Further Research 246
   Summary of Chapter 5 252

6. DEMOGRAPHIC DIMENSIONS OF URBAN GROWTH 255
   Economic Change Under Conditions of Population Growth 256
   Migration 263
   Urban Size 273
   The Patterns of Growth of Urban Areas . . 292
   Systems of Urban Density 298
   The Age-Sex Distribution 312
   Some Studies of Demographic and Urban Change 328
   A Comparison of the Three Theories 363
   Summary of Dissertation 385

REFERENCES 388

APPENDIX A. INDUSTRY CLASSIFICATION 405

APPENDIX B. DESCRIPTION OF URBAN AREAS 408

APPENDIX C. INDEPENDENT VARIABLES AND DATA SOURCES 411

LIST OF TABLES

Table	Description	Page
1	Variables for Which Data are Collected	120
2	Simple Regression of 1961 Variables on Corresponding 1951 Variables	122
3	Simple Correlation Matrix of Eleven Variables Used in Regression Equations	128
4	Regression Coefficients of Six Variables and Contributions at Each Step to Reduction of SS of 1961 Wage Variance in Stepwise Regression	132
5	Standardized Regression Coefficients of Six Variables at Each Step of Stepwise Regression Analysis	133
6	Measures of Importance of Ten Variables in the Regression Equation	138
7	Regression Coefficients of Nine Variables and Contribution of Each Step to Reduction of SS of 1961 Wage Variance in Stepwise Regression	140
8	Standardized Regression Coefficients of Nine Variables at Each Step of Stepwise Regression Analysis	141
9	Measures of Importance of Nine Variables in the Regression Equation	143
10	Regression Coefficients of Six Variables and Contributions of Each Step to Reduction of SS of Change in Wage in Stepwise Regression	146
11	Standardized Regression Coefficients of Six Variables at Each Step of Stepwise Regression Analysis	147
12	Measures of Importance of Nine Variables in the Regression Equation	149
13	Measures of Importance of Variables From Each Theory	151

LIST OF TABLES

Table	Description	Page
14	Measures of Importance of Variables from Sector Theory and Growth Pole Theory	153
15	Simple Correlation Matrix of Fifteen Variables Used in Regression Equations	165
16	Regression Coefficients of Seven Variables and Contributions at Each Step to Reduction of SS of 1961 Wage Variance in the Stepwise Regression	166
17	Standardized Regression Coefficients of Seven Variables at Each Step of the Stepwise Regression Analysis	167
18	Measures of Importance of Five Variables from Each of Growth Pole Theory and Sector Theory	170
19	Regression Coefficients of Eight Variables and Contributions at Each Step to Reduction of SS of Migration Impact Variance in Stepwise Regression	173
20	Standardized Regression Coefficients of Eight Variables at Each Step of Stepwise Regression Analysis	174
21	Average Earnings of Male Wage Earners, 1951 and 1961, and Change in Rank, for Canadian Cities	185
22	Correlation Coefficients of Major Independent Variables with Three Dependent Variables	188
23	Percentage Change in Labour Force Size Attributed to Migration	209
24	Residual Change in Average Wage of Male Wage Earners in Canadian Cities, 1951-1961	211
25	King's Canadian City Groupings, 1951	220
26	King's Canadian City Groupings, 1961	228

LIST OF TABLES

Table	Description	Page
27	Cities Classified by Dominant Function	238
28	The Central Place Functions of Financial Institutions in Twenty Canadian Metropolitan Areas, 1961	245
29	Coefficients of Correlation Among Various Specialisation/Diversification Indices	280
30	Coefficients of Correlation for City Size and Specialisation/Diversification Indices	281
31	Age-Sex Distribution Statistics of Population of Census Metropolitan Areas, 1961	325
32	List of Fifty-Seven Variables Factor Analyzed by Moser and Scott	330
33	List of Variables Used by Mattila and Thompson in Multiple Regression Models of Urban Development	344

Chapter 1

Introduction

The purpose of this chapter is to introduce the proposal that the determination of differential urban growth is an essential component of a broad theory of urbanization. The process of urbanization has been, and still is, one of the most significant in man's history. Any comprehensive theory of urbanization must necessarily explain the causes of urban growth at both the macro and micro levels. One approach at the micro level would include a theory which explains differential urban growth. It would describe the factors that cause one city to grow faster than another.

The chapter begins with a discussion of the meaning of the term 'urbanization'. The meaning has been under debate for a considerable time by those who see it primarily as an ecological concept and by those who regard it as a broader, sociological concept. Both parties recognize that the increase in urban populations vis-à-vis rural populations has played a crucial role in the development of present human society. Thus a valid focus of social science is the determination of the causes of urban growth.

A number of theories have been proposed which deal with this topic but few have focused on the factors which determine which cities grow faster than others. These theories are summarized and the need for an explanation of differential urban growth is emphasized. The suggestion is made that theories of regional development could fruitfully be adapted and applied to urban development. Finally the content of the remaining chapters is outlined.

Urbanization

Urbanization is "without a doubt the greatest social revolution in human history" (Reissman, 1964, p. 170). As such it deserves and has received considerable attention from philosophers and social scientists. It has been defined in a variety of ways, two of which predominate. One can be described as an ecological definition involving the concentration of peoples in built up places and/or the proliferation of these places. It has been defined as a process of population concentration which proceeds in two ways: a multiplication of points of concentration and an increase in the size of individual concentrations (Tisdale, 1942, p. 311). A similar definition makes it "the proportion of the total population concentrated in urban settlements, or else a rise in this proportion" (Davis, 1965, p. 41).

The alternate definition is supported by Louis Wirth who objected to the limited conceptualization of urbanization held by the ecological school. He argued that the crucial aspect was not ecological but socio-psychological and that the ecological approach was detrimentally biasing an adequate conceptualization of the total social process (Wirth, 1938). According to Wirth, the density and size of urban areas affected social relations in a number of ways. In the first place they led to heterogeneity, since an increase in the number of people in a community led to an increase in the possible range of ideas, cultural patterns, occupations, and personal traits.

Another consequence of size and density suggested by Wirth was that one person could not know as many others intensively as in the folk society. As the number of potential interactors increased, the intensity of the interactions decreased. People would come to see only part of the personality of others - to see them in segments of roles. Social relations would become more formal, often impersonal, transitory and segmental.

A supporting point of view is that urbanization is "social change on a vast scale" (Reissman, 1964, p. 154). It is a total process which has irrevocably changed values, roles, and institutions. Cox eloquently puts it this way:

> Urbanization is not just a quantitative term. It does not refer to population size or density, to geographic extent or to a particular form of government... Urbanization means a structure of common life in which diversity and the disintegration of tradition are paramount. It means

> a type of impersonality in which functional relationships multiply. It means that a degree of tolerance and anonymity replace traditional moral sanctions and long-term acquaintanceships (Cox, 1965, p. 4).

Tisdale's rejoinder to the inclusion of a socio-psychological dimension in the concept of urbanization is that "the one trait which can be demonstrated to have a complete and positive association with urban growth is population concentration. If this is the only trait, it should itself be the definition" (Tisdale, 1942, p. 312). But Reissman argues that an adequate theory should be broad enough to account for the essential features of the urban environment (Reissman, 1964, p. 122). If life styles and values are in fact central features of urbanization, they should be included in the definition of the concept. He attempts to contribute in this direction by proposing a four-fold framework of components of urbanization (Reissman, 1964, p. 168). These are:

1. urban growth;
2. industrialization;
3. emergence of a middle class into power;
4. rise of nationalism as the dominant and political ideology.

In other words, urbanization is a process which encompasses increasing proportions of urban dwellers, the shifting of economic activities from agriculture to manufacturing, restructuring of power relations resulting in a rise in a middle class which stood to benefit most from industrialization, and a rise in an ideology which justified changes, commanded loyalties, and motivated actions.

Writers who prefer the broad definition of urbanization concede that urban growth is a central, if not the crucial, aspect of the broad process they call urbanization. The point of this section is not to resolve the debate but to show that whatever one's analytical framework, the part played by urban growth is crucial in human social processes.

Urban Growth

This section attempts to reflect the magnitude of the scale with which world and Canadian populations are undergoing urban growth. Cities are a relatively recent phenomena in man's history. They first appeared some 5,500 years ago but their existence at that time was precarious (Davis, 1965, p. 41). At the time of the birth of Christ there were approximately twelve cities of 100,000 population in the Mediterranean area, and perhaps a few others of that magnitude in Asia (Hawley, 1950, p. 371). By 1800 there were only twenty-one cities of 100,000 or more in the world (Hawley, 1950, p. 372). But there were 146 such cities by 1900, 202 by 1920, 537 by 1927, 720 by 1940, and 1300 by 1960 (Hawley, 1950, p. 372; Davis, 1972, p. 18).

Hauser traces the growth of the urban population in the world from 1800 to 1950 for cities of 100,000 or more, for cities of 20,000 or more, and for urban places of 5,000 or more. Over that time span the world population increased two and one-half times, while the population of cities of over 100,000 increased 20 times, the population of cities of over 20,000 increased 23 times, and that of urban places of 5,000 or more increased 26 times. Over the same time span the percentage of the world's population living in cities of 100,000 or more increased from 1.7 to 13, the percentage living in cities of 20,000 or more increased from 2.4 to 21, while the percentage living in urban places of 5,000 or more increased from 3 to 30 (Hauser, 1965, p. 7).

Since 1950 the rate of increase of urban residence has been twice that of the preceding fifty years (Davis, 1965, p. 41). Davis estimates that the percentage of the world's population living in urban places increased from 28.2 in 1950 to 38.6 in 1970. Over the two decades from 1950 to 1970 the percentage of the world population living in cities of 20,000 or more increased from 22.7 to 32.2 while the percentage living in cities of 100,000 or more increased from 16.2 to 23.8 (Davis, 1972, p. 51).

We may note some discrepancies between Davis' and Hauser's figures. Apparently they used different sources for compiling their percentages. However the incredible aspect of these figures is that they are percentages. As the base, which is the total world population, increases, it becomes more difficult for percentage figures to show large increases. Yet world urbanization figures, expressed as percentages, continue to increase at increasing rates.

Canada, in spite of its late start as a populous area and in spite of its early function as a colony and provider of raw materials, was always one of the most urbanized countries in the world (Stone, 1967, p. 10). In 1825, the cities of Montreal and Quebec, with populations of 32,000 and 22,000 respectively, comprised five percent of the populations of Upper and Lower Canada and the Maritimes (Stone, 1967, p. 14). This is twice the percentage in cities of 20,000 or more for the world in 1800 as estimated by Hauser. The percentage of the Canadian population living in urban places was 13 in 1851, 18 in 1871, 30 in 1891, 42 in 1911, 53 in 1931, 62 in 1951 and 76 in 1971 (Stone, 1967, p. 29; Statistics Canada, 1973, Table 10). These figures indicate that urban growth is a significant force in Canadian history and that its causes and consequences are worthy of thorough analysis.

Causes of Urban Growth

An important question that is often raised deals with the original factor or factors in the emergence of cities and the causes of their continued growth. This section briefly summarizes the major theories that have developed regarding these factors.

Urbanization and urban growth have come under a considerable amount of scrutiny by social scientists in North America and Europe since the middle of the nineteenth century. North American social scientists have concentrated on two aspects of urbanization (Martindale, 1958, p. 46). They have been guided by ecological theory aimed at discovering laws which determine the physical shape of the city and by socio-psychological theories which are concerned with differences in social and psychological traits manifested by rural and urban dwellers. The socio-psychological theories were generated by a reaction to the absence of social variables in ecological theories.

European theorists were more concerned with the broader aspects of the city as part of human society. Two general properties of the European approach to the study of the city are, first, that characteristics of any unit of social life are determined by institutions, and second, that the city is an evolutionary or historical product (Martindale, 1958, p. 46).

One of the earliest European theorists, Fustel de Colanges, maintained in 1864 that the crucial institution which led to the development of the city was religion

(Martindale, 1958, p. 46). He points out that originally the family was the only institution in society. A primitive religion constituted the family and dictated its laws. The same religion was gradually enlarged to encompass several families living in one place. A citizen of a city was then defined as one belonging to the religion of that city.

Keutgen in 1895 and Maitland in 1898 argued that the city was originally a military institution affording a stronghold where the inhabitants surrounding the place could retreat for protection (Martindale, 1958, p. 50). Their explanation has gained considerable popularity although subsequent theories expanded to include concomittant factors. Pirenne believed that walled fortresses were only the places of origin of cities. Cities were formed by economic institutions. They came into being as the result of the activities of merchant caravans which settled outside the walls. The caravans introduced a new class of merchants who brought about new laws, private jurisdiction, and free property (Martindale, 1958, p. 50). For Pirenne the city was a community of merchants.

The origins of the city were credited to institutions such as the family, religion, the military, and economics. These causes were so selected because of the European bias towards institutional explanations. Another group of theorists were not so much concerned with the origins of cities as they were with the growth of cities. One of the most comprehensive studies of this kind was carried out by Max Weber.

Weber called the medieval city a marketplace. "The local market forms the economic center of the colony in which, due to the specialization of economic products, both the non-urban population and urbanites satisfy their wants for articles of trade and commerce" (Weber, 1958, p. 67). However his concept of the city extended beyond this simple model. The economic dimension of the city was a necessary but not a sufficient condition for its growth. He recognized a number of other conditions that varied in time and place so that no one factor could be singled out as the sole source of city growth. But he emphasized the military institution as a primary factor in the emergence of cities in antiquity and the economic institution as a primary source of the emergence of medieval cities (Weber, 1958, p. 212).

Two of the first Americans to attempt to analyze city growth were Charles Cooly in 1894 and Adna Weber in 1899. Cooley's work, a forerunner of later ecological studies to proliferate in America, concentrated on factors affecting the location of new towns. It is dealt with in later paragraphs.

Adna Weber saw urban growth primarily as a product of a number of economic forces generated by the industrial revolution (Weber, 1963, p. 156). These include the emergence of transportation systems and means of communication to commerce and commercial centres, improvement in agricultural methods, mercantilism, and the emergence of the factory system (Weber, 1963, pp. 160-196). He also allowed for a number of secondary or individual causes of migration to cities. He classifies these as economic, political, and social factors of urban growth. The secondary economic factors include higher wages and more varied job opportunities in cities (Weber, 1963, pp. 210-213). The secondary political causes include legislation promoting freedom of trade and migration. Free trade favours the growth of cities by enlarging markets while migration provides the city with a source of labour from the rural population. A third political cause is centralized administration which provides more job opportunities in large cities while a fourth is land tenure which prohibits ownership of land by many rural inhabitants (Weber, 1963, pp. 213-217). The social causes include better educational and recreational facilities, a higher standard of living, and greater opportunities for intellectual association (Weber, 1963, pp. 218-220).

While Weber addressed the question of what forces in human society caused the growth of cities, Cooley turned to a complementary question: why are cities located where they are? He developed a theory of transportation in his Ph.D. thesis which attributed a break in transportation as the primary reason for a city's location (Martindale, 1958, p. 16). The break may occur at a river crossing, at the junction of a railway and a river, or near an ocean harbour. The transportation break required considerable equipment and facilities to move goods from one carrier to another, resulting in the aggregation of capital and labour at that point.

This theory has been retained in the present, prevalent threefold classification of reasons for locations of cities (Berry, 1964a,p. 121):

1. strategic locations on transportation routes (this is Cooley's theory of transportation);
2. the outcome of local concentrations of specialized economic activities such as mining towns and administrative centres;
3. central places performing retail and service functions for surrounding areas.

Sjoberg brought into relief the dichotomy of the preindustrial and the industrial city. Although each developed as a function of a different set of forces, the rise of technology and a changing social organization within the setting of a favourable environment played a crucial role in the growth of cities before and after the industrial revolution (Sjoberg, 1965, p. 55). Factors leading to growth of the preindustrial city were literacy, specialization of labour and improved methods of production. The growth of the industrial city was based on the harnessing of inanimate energy and the fluidity of the class structure.

Sjoberg's factors, technology, social organization, and environment, constitute three of the four components of the framework known as the ecological complex. Sjoberg may have omitted the fourth, population, because it is obviously a contributing factor to urban growth and he did not see a need for its elaboration. Hauser adds this fourth component to Sjoberg's tripartite framework to complete the ecological complex as a method of classifying factors of urban growth. He states that the emergence and development of the city was necessarily a function of (Hauser, 1965, p. 1):

1. the size of the total population;
2. the control of the natural environment;
3. technological development;
4. development in social organization.

He argues that the first two are obvious components because large aggregates of humans necessarily require a large population plus an environment amenable to their maintenance. At the same time agricultural technology had to achieve a level which made surplus production of food possible. This freed persons to participate in non-agricultural activities.

Finally an organization had to emerge which allowed cooperation within the population aggregate and between it and its hinterland.

A recent theory of urban growth which has not gained much popularity is Meier's communication theory of urban growth (1962). Meier invokes an entirely new set of concepts, borrowed from cybernetics and information theory, to provide a framework for analyzing factors, processes, and limits of urban growth. His theory is based on the claim that "cities were evolved primarily for the facilitation of human communication" (Meier, 1962, p. 13). Cities have grown because of the evolution of communication technology from face-to-face intercourse to notched sticks, clay tablets, musical instruments, writing, the printing press, telegraph, telephone, radio, television, and computer. He argues that further development of cities requires that the process of collection and interpretation of information from the environment must supply new concepts and images more rapidly than they are lost through attrition, obsolescence, and interference (Meir, 1962, p. 12).

Three principles of information theory are applied to urban growth (Meier, 1962, pp. 150-152). The first principle is that "if a society of mortal individuals is to survive, information must be conserved." Otherwise the society loses its organization and structure. A second principle states "a sector of society that grows in influence, wealth or power measured in absolute terms, must experience a growth in information flow that occurred prior to or simultaneously with the other recorded growth." A condition is imposed here: as people approach their channel capacities for communication it is possible that the inadvertant costs of errors will reduce wealth, influence or power. Because of this condition, the third principle states that "if advanced societies are to increase their organization (and capacities for cultural interaction), the interactions between automata in their service must increase even more rapidly."

Meier's communication theory of urban growth is less a theory than a set of borrowed concepts providing a radical new framework for studying urban growth. However the concepts have not been tied together systematically nor have they been operationalized. Until this is done, the theory remains fairly sterile.

Up until now, this section has summarized studies dealing with three questions: what is the origin of cities, what are the factors of urban growth, and what are the reasons for specific locations of cities? A question related to the above three and central to the present thesis is: why do certain cities grow faster than others? Only two economic theories which deal with this problem have received widespread attention: export base theory and central place theory (Richardson, 1971, p. 82). No more will be said of export base theory at this point as it receives considerable elaboration in Chapters Two and Three of the thesis.

Central place theory was laid out by Walter Christaller in 1933 and subsequently elaborated by August Losch in 1940. The six main features of Christaller's theory as summarized by Berry and Pred (1965, pp. 3-4) are as follows:

1. the chief function of a town is to be the centre of its rural environ and mediator of local commerce with the outside world; therefore, cities are located central to the maximum profit area they can command;
2. within a system of towns, the greater a town's centrality, the higher its order;
3. Higher order places offer a larger range of goods and services, but are more widely spaced than lower order places;
4. lower order places offer goods purchased most frequently;
5. a hierarchy of central places exists;
6. three principles are developed in order to explain the number, size, and distribution of central places: a market principle, a transportation principle, and an administrative principle.

According to the market principle, central places lie at certain equal distances and in proper orientation with respect to central places of higher order. They are grouped so that there is a certain number of places belonging to each order. Two essential concepts used to explain the market principle are 'range' and 'threshold' of a good. The range of a good is the maximum distance a person will travel for that good. The threshold is the minimum amount of purchasing power necessary to support the supply of a good from a central place (Ray, 1968, p. 48). Goods with larger thresholds and ranges are located in higher order places.

The transportation principle dictates that central places lie along lines running between central places of a higher order. The administrative principle recognizes the influence of political boundaries on the distribution of central places. The market principle is the dominant of the three while the other two explain deviations from the ideal, hexagonal pattern.

Lösch attempts to justify Christaller's hexagonal market areas as the most advantageous shape for economic regions in terms of both supply and demand (Lösch, 1954, pp. 105-114). Ideally the market area is circular with a central point of production. But competition from producers outside this trading circumference compels those within to contract their intersecting circles into market areas of hexagonal shape. To order the landscape systematically, Lösch groups his hexagonal nets for particular products and services by the size of their respective market areas. The size distribution of cities thus depends on how many particular goods and services of what size markets coincide at any individual central place (Lampard, 1968, p. 88).

According to central place theory, differential growth occurs because higher order places offer a larger range of goods and services than lower order places as the result of increasing specialization and the break-down of self-sufficiency brought on by industrialization, commercialization, and technology. This causes greater interdependence and consequently greater competition among central places. However, the question of why certain central places attain a higher order than others is not explained. Presumably they do so because they are more central, due to an abstract geometric law that predetermines a place's order.

A false assumption of central place theory is that a city's growth is a function of the size of its hinterland population. Cities grow for reasons other than servicing their hinterlands and large cities do not necessarily specialize in higher order goods and services (Richardson, 1971, p. 82). Although central place theory can provide a useful framework for analyzing an urban system, particularly a young urban system in a large rural region, it does not promote an understanding of differential urban growth in present-day industrial countries.

The intent of this chapter is to reveal that no adequate theory of differential urban growth exists. Such a theory would be a useful component of a theory of urbanization, a process which is presently playing a major role in the restructuring of human society. A useful source from which the beginnings of a theory of differential urban growth can be drawn is the field of regional studies. Regional science has been extensively concerned with the problem of growth within regions. It is a very short step from the study of growth of a region to that of growth of a city, particularly in industrial parts of the world where the bulk of economic activities are non-agricultural.

Within the North American context, the concept of a region first gained momentum in the southern United States in the 1930's. It evolved as the result of efforts to overcome the effects of the depression. It was defined in terms of a physically homogeneous area held together by a distinctive folk culture (Friedmann, 1956, p. 3). As the concept gained acceptance in the northern United States, and as urbanization progressed, the concept evolved to become a physically heterogeneous area dominated by a city.

Regional science has seen the construction of three major theories of regional development: sector theory, export base theory, and growth pole theory (Perloff et al., 1960, pp. 57-60). Sector theory states that a region develops as the major components of its economy shift from primary to secondary and from secondary to tertiary industries. Export base theory states that a region develops because it has a commodity which is in demand outside the region. Growth pole theory states that a region develops because industries which use each other's commodities are clustered in one place achieving economies of agglomeration.

The purpose of this study is to test the applicability of these three theories to differential urban development in Canada from 1951 to 1961. The major problem can be phrased as following: which theory of regional development best explains urban growth in Canada during the period 1951 to 1961? The acceptance of one theory does not preclude the acceptance of others. Perhaps all three are necessary for explaining urban growth in Canada during the decade.

Outline of the Remainder of the Thesis

The outline of the remainder of the thesis is as follows. Chapter Two provides descriptions of the three theories and summarizes research in the literature aimed at testing the three theories. The chapter also provides a comparison of crucial similarities and differences among the three theories.

Chapter Three contains a discussion of the methodology of selecting the most effective operational definitions or indicators of crucial variables in each theory. Included in Chapter Four is a description of the steps involved in the collection, processing, and analysis of the data.

Chapter Five provides a discussion of the findings and a comparison of these with other research findings of Canadian urban growth for the decade 1951 to 1961. An attempt is made to synthesize these findings with prevalent ideas in sociology, demography, and economics concerning urban development.

Finally, Chapter Six provides an analysis of the demographic and economic implication of urban growth and development. The three theories of regional development are compared in terms of their implicit consideration of the demographic and ecological dimensions of growth. The chapter ends with a summary of the thesis.

Chapter 2

Theories

This chapter describes the three theories to be compared. A brief outline of the main points of each theory is followed by an elucidation of the details of the theory, a summary of the research done to test hypotheses derived from it, and mention of major criticisms of the theory found in the literature. The chapter ends with a brief comparison of the three theories.

Export Base Theory

Douglass North has put forward a five-stage regional export base model that he suggests could be applied to most market-exchange economies with no population pressure problems:

1. A brief subsistence stage.
2. Rapid development in the exporting of staple commodities to more advanced regions as the basis of the regional economy.
3. With the growth of external economies, inflows of capital, and provision of an export oriented infrastructure, a stage of export intensification and regional development occurs.

4. In time, residentiary or locally based industry develops to serve the local market.
5. Finally, the expansion of resident industries, together with footloose industries located by chance, may reach a point where they too enter the export market, thus diversifying the region's export base (Hurst, 1972, p. 330).

Export base theory derives from Keynesian income theory applied to an open economy and adapted to a long run analysis (Richardson, 1969, p. 336). The gist of Keynesian income theory is that investment is the prime factor in raising the level of national income. Investment will occur only when demand for a commodity rises, guaranteeing returns on investment in that sector (Siebert, 1969, p. 24). National income is increased not only by the amount of investment but also by the resulting increase in income in other sectors generated by the original investment - called the 'investment multiplier'.

Investment is the addition to the existing stock of real capital assets; e.g., construction of factories, office buildings, and transportation facilities. Keynes reasoned that new investment led to increased consumption which in turn resulted in increased income (Dillard, 1948, p. 87). The investment multiplier is the ratio of the change in income and the change in investment and is a function of the propensity to consume which is the proportion of additional income spent on consumption. Keynes originally intended the multiplier to refer to the effect of investment in the capital-goods industries upon employment in other sectors of the economy and vice versa - the effect of investment outside capital-goods industries upon employment in that industry. The concept has been broadened to refer to the effect of investment in basic industries on regional and urban growth (Pred, 1966, p. 30).

Keynes' theory was drawn up for countries which are classified as closed economies but a region must be conceived as an open economy because of the frequency with which it exchanges commodities and services with other regions.

Three aspects of the frame of reference of export base theory should be emphasized. The theory is applicable only under conditions of factor mobility in a capitalist system. Furthermore the region must begin with a low population density and must have resource endowment or production possibilities.

Given the above conditions a region has a possibility of growing according to the process described by export base theory. In the initial stage natural endowments will dictate what the export commodities are (North, 1961, p. 3). The growth of the region may be a function of the type of resource which initiated the growth. A region's resource endowment is 'good' to the extent that production of the resource (1) expands, (2) encourages locationally-associated forward linkages, and (3) guarantees a high regional multiplier (Perloff and Wingo, 1964, p. 226).

The extent to which production of a commodity can expand is a function of two conditions: the supply and demand conditions of the country with respect to that commodity, and the quantity of the resource endowment. The latter point is self-explanatory - a limited supply of the resource does not encourage a large amount of investment in its extraction and distribution.

Conditions of supply and demand in the country dictate the amount of investment expended in distributing the resource throughout the country. Location of the same resource in alternate regions reduces demand for the commodity in the subject region and increases competition for the supply of the resource. Assuming that the subject region has a competitive advantage over others in producing the commodity, an essential condition for increased production is that there be a high income-elasticity of demand for the resource (Perloff and Wingo, 1964, p. 226).

The second necessary condition for growth is that the resource endowment must have location-associated forward linkage effects. This means that industries which use the resource as inputs will be attracted to the region to reduce transportation costs. Location of new firms in the region leads to more intensive use of labour and capital and stimulates the formation of social overhead capital such as transportation routes and energy sources (Hurst, 1972, p. 329).

The third condition, a multiplier effect, occurs when the growth of an export industry has an effect on the growth of other interdependent industries. The iron and steel industry is an excellent example of one with high multiplier effects because most of its output is used as input to manufacturing industries. Agricultural land is generally considered to be a relatively poor resource for future development because agricultural products go directly to the consumer allowing little opportunity for forward linkages. Agricultural produce has a low income-elasticity of demand, compounding its inadequacy as an export base. Most minerals are excellent resources with high linkage and multiplier effects but have the disadvantages of being non-renewable and vulnerable to substitutes (Perloff and Wingo, 1964, p. 226). Oil and gas are usually poor base industries because much of the output goes directly to the consumer, reducing opportunities for linkage and multiplier effects. The oil and gas industry is capital intensive causing most of the profits to go to capital which when imported is lost to the region where the oil and gas is found until recently.

If capital is from the region, of course, oil and gas represent valuable resources to the region.

The prerequisites for growth according to export base theory include a resource endowment which is in demand outside the region and which can be produced and transported out of the region at a lower cost than can be done by competing regions. The resource must have a high elasticity of demand and strong linkage and multiplier effects. Given these conditions the region has an opportunity to grow. The subsequent steps necessary for continued growth are (1) a development of nonbasic industries, and (2) a broadening of the export base (Tattersall, 1962, p. 217). Firms must be attracted to the region to provide services and domestic commodities to the basic industries. Consequently, export base theory must encompass location theory to explain the attraction of nonbasic industries.

A brief overview of location theory will familiarize us with some of its basic concepts. It assumes that a firm will choose that location which allows it to maximize profits. Another assumption usually held is that costs of capital, labour and energy as well as depreciation and interest rates are equal throughout all regions. If this is not so, they must be considered part of a firm's decision to locate. If we can hold these factors constant, the two general factors determining a firm's location are transfer costs and agglomeration economies.

Transfer costs are those costs necessary to transport inputs and outputs to and from the location of the firm. They include all transportation, storage, and communication expenses. If an industrial process involves considerable loss of weight, the firm will tend to locate close to the source of its inputs to reduce input transfer costs. If transfer costs of outputs are higher than those of inputs, the firm will attempt to locate close to a market. Alfred Weber judged transfer costs to be the most fundamental determinants of the location of a firm (1929, p. xxvi).

Agglomeration refers to the massing together of firms for reasons of linkage effects and external economies. The linkages may be forward, backward or complementary. Forward linkage effects refer to the effects of outputs on consumers and is of concern to primary industries. They will benefit if their consumers increase consumption or move closer to the producing firm to reduce transfer costs. Backward linkages are those linkages which influence the firms supplying inputs to an industry. Manufacturing industries for example are concerned with influencing their suppliers to reduce cost prices and improve supply. Complementary linkages occur when would-be competitors benefit by their mutual proximity. Such benefits occur when buyers and sellers are attracted to the area by the total demand or supply of several competing firms. Complementary linkages reduce transport costs because they allow larger volume shipments. They also improve the likelihood that adequate service and storage facilities will be available. Examples of this linkage effect are in every large city where shoppers go to a specific part of the city because they know they can find a good selection of a particular commodity such as shoes or ladies' wear.

The other aspect of agglomeration, external economies, includes all advantages occurring for reasons other than increased division of labour within the plant or improved technological and organizational processes. These are called internal economies. External economies include advantages accruing from the availability of social overhead such as electricity, water, and storage and transport facilities. They may also be gained from close personal contacts with business associates, and from the existence of a larger pool of skilled labour. If the higher densities of industries and personnel results in higher rather than lower costs to the firm, it is called deglomeration. Often the coincidental agglomeration of a number of independent industrial complexes results in high rents, pollution, and overcrowding of transportation facilities characteristic of larger cities. These represent deglomerative factors for each individual firm in spite of the fact that they derive from the pursuit of advantages of agglomeration (Isard, 1956, p. 139).

Transfer costs and the effects of external economies and linkages determine which alternate region a firm will select for location. In the early stage of growth forward-linked industries will be attracted to an exporting region because of demands for such industries caused by increasing returns to the staple industry (North, 1955, p. 253). These are linkage effects. These industries serve to broaden the base and reduce costs of production. Tattersall believes that a broad base is necessary for sustained growth because a single industry will experience growth retardation and because productivity increases are more likely in new industries (Tattersall, 1962, p. 233). The resulting agglomeration of base industries creates a market for non-basic industries. Some of these will expand and may join the basic sector by exporting some commodities and services (North, 1955, p. 255).

The final step in export base theory is industrialization which North defines as an increasing proportion of the export base being made up of finished consumers' goods and finished manufactured producers' goods (North, 1955, p. 254). Most manufacturing industries are market-oriented, causing this step to be the final one in the development process (Tattersall, 1962, p. 217). It only occurs after enough agglomeration has preceded to ensure a market for manufactured goods within the region itself. Other conditions necessary for industrialization are:

1) efficient nonbasic industries that reduce social overhead costs;
2) major improvements in transportation;
3) growth in income and demand in other regions;
4) technological developments that reduce processing costs (North, 1955, p. 255).

Furthermore at any point in time an exporting region may suffer a decline because of:

1) a decline in demand outside the region;
2) exhaustion of the natural resources;
3) increasing costs of land or labour relative to those of competing regions;
4) technological changes that alter the relative composition of inputs (North, 1955, p. 254).

Thus we see that regions with resource endowments may achieve export expansion but will not continue to grow unless they attract nonbasic industries. Upon achieving this, they must undergo the final necessary condition for continued growth, which is industrial diversification (Tattersall, 1962, p. 218). Even then they are subject to exogenous forces which may retard growth.

Research

Most research on export base theory has been carried out in the United States (Keeble, 1967, p. 267). A thorough investigation was done by Borts who tested two alternative explanations of regional growth (1960). One explanation was export base theory which implies that a growing region's wages will increase relative to other regions' because of advantages in production functions, transportation costs, and marginal efficiency of investment. The alternative theory was that regions grow because of differences in initial endowment of capital and labour. This condition results in earning differentials which cause interregional movements of capital and labour resulting in a convergence of wages among regions. Borts examined variations of regional wages over several periods in the United States and found that wage divergence occurred more often than convergence, supporting export base theory.

Tattersall investigated the development of the United States Pacific Northwest from 1880 to 1960 (1962). He noted a rapid growth of the region from 1880 to 1910 when transcontinental railroads linked it to the rest of the country. The primary exports during that period were wheat and lumber, both of which had an elastic long-run supply. From 1900 to 1920 there was a rapid growth of nonbasic industries which appeared to be tied to the export sector. However regional growth slowed from 1910 to 1940 because of a reduced demand for wheat and lumber. Tattersall points out that the slowdown might not have happened if the export base had been broadened.

World War Two brought impetus for further growth principally because of the location of an aircraft industry which provided new exports from the region and because of an increased demand for the original staples. But growth lagged shortly after the war ended. Tattersall believes that at this time the region should have industrialized but did not because of (1) the limited size of the regional market, (2) a lack of supply of cheap fuel, and (3) modest intraregional linkage effects. He endorses export base theory as the best explanation of growth in the Pacific Northwest.

Perloff et al. (1960), Perloff and Wingo (1964) and North (1961) look at the historical growth of the United States in terms of the interdependent development of its various regions. They agree that the responsibility for American growth rests primarily on the exports of first, agriculture in the midwest, second, minerals in the mountain regions, and finally services from industrialized regions.

Criticisms

A number of criticisms have been directed at export base theory. One of the most serious is that it attempts to explain growth in terms of an exogenous variable, demand, which itself is not explained. As a theory it should retain the most important variable within its explanatory framework (Siebert, 1969, p. 98). Another contention is that the single variable, external demand, is over-emphasized. Tiebout insists that the prime mover is regional income which is a function not only of demand but also of national government expenditures and business investments in the region (1956, p. 161). North admitted that these would indeed affect a region's growth but argued that the only necessary variable to be considered was external demand if government expenditures could be recognized as a qualification to the theory (1956, p. 166).

Another criticism related to the overemphasis of external demand is that nonbasic activities also play a crucial role in the region's development and should be weighted as heavily as basic industries. Tiebout argues that exporting industries are attracted to a region and thrive after locating there because the nonbasic industries are providing more efficient services than their counterparts in other regions (Tiebout, 1956, p. 163). Blumenfield argues that metropolitan areas survive or grow because they contain:

1. large and variously-skilled labour forces which are present because of local consumer services such as housing, schools, retail units and local transportation facilities;
2. business services such as power sources and interurban transportation and communication systems;
3. local and regional markets (Blumenfield, 1955, p. 131).

Consequently exporting industries depend on these services making service industries 'basic' and exporting industries 'nonbasic'. A similar argument has been made that internal as well as external demand should be integrated in the model as a factor of growth (Perloff et al., 1960, p. 60).

Siebert presents several criticisms but he seems to equate export base theory with the single article written by North in 1955 (Siebert, 1969, pp. 96-100). Several of his comments allude to points North omitted such as the obvious fact that the theory assumes potential resources are not fully utilized. However he also makes some very relevant criticisms. Besides expounding on the problem that demand is crucial but exogenous in the theory, Siebert shows that the theory does not discriminate between the export of investment goods and consumption goods. Only the latter should be considered because only they add to a region's welfare.

Another criticism by Siebert is that export base theory fails to recognize terms of trade as a factor of growth. Terms of trade is defined as the relative price of import and export commodities. It is the ratio of the price index of all export goods and the price index of all import goods. Variation in terms of trade may reduce income in a region even though determinants of growth have expanded. Blumenfield is critical of the basic-nonbasic approach, bringing into his argument examples such as subsistence Indian villages which did not trade with other villages, and medieval towns which survived by taxing peasants rather than by contributing to the national economy (Blumenfield, 1955, p. 117). However, he forgets that these fall outside the domain of export base theory which is applicable only to market-exchange economies.

Blumenfield disagrees with the distinction between basic and nonbasic goods as those which bring money into the community from outside sources and those which do not. He suggests that the distinction should be between those industries which "by the nature of their product have to and can compete with outside producers, regardless of the location of their actual sales, and those which do not compete" (Blumenfield, 1955, p. 119) [italics in the original]. He provides an example of a General Motors plant in Flint, Michigan which, if shut down, would cause several department stores to close. But the closing down of a department store would not force the General Motors plant to close. He states that the plant is more critical to the local economy because it must compete with automobile plants outside the region while department stores compete only with other stores within the region (Blumenfield, 1955, p. 121). If a department store closes, the income of other competing stores would gain by what that store lost, but if the General Motors plant closes down, there is no alternative source of income to compensate for the region's loss of income provided by the automobile plant.

Export base theory has suffered from telling criticisms. As a theory it has the weakness of using an exogenous variable to explain changes in all endogenous variables. It places too much weight on external demand to the exclusion of such factors as terms of trade and internal dynamics of nonbasic activities. Richardson suggests that in spite of its theoretical inferiority, export base theory is useful as a planning and forecasting tool provided its limitations are kept in mind (1969, p. 254).

Sector Theory

North describes sector theory as a process of five steps:

1. The region has a self-sufficient subsistence economy. The population is scattered throughout the region according to the location of resources.
2. Improvements in transportation result in some trade of primary products with other regions. This leads to specialization of agricultural activity and attracts various service industries.
3. As interregional trade increases, there is cultivation of crops which involve more intensive land use, greater profits and more skills and technology.
4. With increased agricultural population, per capita returns to agriculture begin to decrease. This results in industrialization. The industries which develop tend to be those which utilize local resources, such as processing of food, manufacturing of wood products and preparation of textile fibers.
5. Finally there is specialization in tertiary industries and export of capital, skilled personnel and services (North, 1955, pp. 244-245).

Sector theory is based on the empirical findings of Clark and Hoffman. Both sifted through available historical data of all developed countries. Clark and Fisher are credited with making the hypothesis that an economy grows as its major production shifts from primary to secondary and from secondary to tertiary industries. Hoffman established a closely related hypothesis that industrial development occurs when the output of capital-goods industries increases proportionately to consumer-goods industries. His hypothesis is easily integrated into stage four of sector theory as outlined by North.

The essential factor leading to growth during the first three stages is reduced transfer costs resulting from an improved transportation system (Hoover, 1948, p. 188). Once these steps have been successfully completed the crucial problems become how and why industrialization is achieved in a region which has hitherto been successful only in agriculture.

The movement from the third to the fourth step is fraught with difficulties. An oft-quoted reason a region does not develop is the 'Malthusian devil'. High birthrates, which are often found in agricultural regions lead to population pressures. These tend to perpetuate poverty, cancelling any effects that investment might have on the region's growth (Hoover, 1948, p. 189). It is necessary that agriculture be specialized at this stage. The resulting extension of knowledge of agricultural and other methods of production provides the area with a social environment conducive to industrialization. Once a specialized agricultural scenario has been established the greatest difficulty is accumulating capital necessary for investment in manufacturing enterprises (Fisher, 1933, p. 380).

Sector theorists argue that the only way a region can sustain growth is by industrialization. The logic behind this is based on Engel's Law which states "The poorer a family, the greater its proportion of total expenditure used for food" (Woytinsky and Woytinsky, 1953, p. 269). Sector theory takes rising per capita income as a necessary condition for regional growth. Thus as per capita income rises families will find more of their expenditure used for the purchase of commodities produced by secondary industries. This reflects the greater income elasticity of demand attributed to manufacturing than to agricultural products.

The fourth stage of sector theory, industrialization, can be divided into three sub-stages:

1. a period when consumer-goods industries dominate; during this period the ratio of output of consumer goods and capital goods is about 5:1;
2. a period during which there is a rapid growth of capital-goods industries relative to growth of consumer-goods industries;
3. a period during which the growth of capital-goods industries levels off relative to that of consumer-goods

industries. The ratio of total output of the two stabilizes at about 1:1 (Spengler, 1949, p. 87).

During the early stages of industrialization the leading industries tend to be food processing, wood products and preparation of textile fibers (Hoover, 1948, p. 192). Such industries are most likely to initiate industrialization in an agricultural region because they are dependent upon products of local primary industries. They supply elementary human needs, making them relevant to the rural population. They are viable in such a region because no unusual skills are required.

The second sub-stage is crucial because it involves the diversion of investments into capital-goods industries. The transition of dominant production from consumer goods to capital goods is considered essential to continued growth for a number of reasons. Böhm-Bowerk first pointed out that greater productivity obtains when more roundabout methods of production are used (Fisher, 1933, p. 381). Elaboration of the production process results in greater specialization, more linkages, and a higher input of capital, all of which may result in greater output per unit input. Capital-goods industries tend to use more roundabout methods than consumer-goods industries. They also are geared to large-scale production, allowing them to benefit from economies of scale (Hoselitz, 1961, p. 539).

The rise in production of capital goods may signify the 'take-off' which Rostow has popularized. He suggests that a region may reach a point where its growth will either stagnate or spurt, depending upon a number of factors (Rostow, 1956, p. 29). Take-off is the expansion of new industries, new processing techniques and, most important, increasing investment. Diverting investment from agricultural and consumer-goods industries to capital-goods industries is difficult because potential investors may have established patterns of saving rather than of investing, or they may be investing in those industries which have previously rewarded them with small but certain returns (Fisher, 1933, p. 387).

The third sub-stage involves a crystallizing of the capital goods-consumer goods ratio at about 1:1. Both sectors continue to grow. This phase represents the culmination of the industrialization process and the emergence of the region into a stage of industrial maturation. It is marked by exploitation of intensively produced domestic materials and processing of imported materials. The region is capable of importing goods and processing them at lower costs than other regions because it possesses superior skills and technology. At this stage we find smelting, refining and processing of metals in the region as well as oil refining and chemical industries based on coal, petroleum, potash, salt and other minerals (Hoover, 1948, p. 193).

Very little has been written about the final stage of growth because proponents of sector theory have been preoccupied with the industrialization process. Only a few regions have managed to make the transition from agriculture to industry but these few have continued the trend by expanding on tertiary activities such as commerce, transportation, finance and administrative, professional and personal services. The transition from manufacturing to services has not appeared to be a problem.

Research

Testing the Clark-Fisher, and to a lesser extent the Hoffman, hypothesis has been fairly extensive. The first specifies that as a region grows its primary activities are usurped by secondary and later tertiary activities. Hoffman's hypothesis is that the essential reason for growth of an industrial country is the increased role played by capital-goods industries relative to consumer-goods industries.

Marczewski studied the economic growth of France from 1660 to 1958. He found that for each unit increase of the industrial rate of growth, the agricultural rate of growth has risen on the average by 0.27 (Marczewski, 1961, p. 379). The entire period was dominated by a continuous increase in manufacturing's share of total production.

Harris found the Clark-Fisher hypothesis inapplicable in New England. Connecticut was the most productive state but had the lowest percentage of tertiary workers. Vermont was least productive but had the highest percentage of its labour force in tertiary industries (Harris, 1952).

Perloff improved on Harris by considering the whole of the United States rather than one region. He observed that in 1950 those states with a high proportion of their workers in tertiary industries and with a low proportion involved in primary industries tended to have higher per capita income (Perloff, 1957, pp. 164-165). However, he also noted considerable variation from this tendency.

Mehtra tested the Clark-Fisher hypothesis by comparing the industrial structure of the urban labour force of Burma and the United States. He rejected the hypothesis because 63.2 percent of the Burmese labour force was engaged in tertiary activities while 60.4 percent of the United States urban labour force was so engaged (Mehtra, 1961, p. 167). This illustrates a difficulty in comparing developed and underdeveloped countries with respect to proportionate labour force activity in service industries. Proportions in service industries tend to take a U-shaped distribution along the continuum from poorly developed to highly developed countries. Both extremes have large proportions of their labour force in service industries compared to other countries.

The following two studies comparing consumer-goods production and capital-goods production support Hoffman's hypothesis. Hoselitz tested the hypothesis across countries at one point in time. He found that countries in higher economic levels had higher ratios of capital goods/consumer goods output (Hoselitz, 1961, p. 541).

Patel looked at rates of growth from 1860 to 1958 for all major industrial countries. He found a continuous decline over time in the ration of consumer goods to total industrial output. All countries followed this pattern in spite of differences in natural resource endowment, accumulation of technical skills, period of industrialization, speed of growth, experience in international trade and capital movement, fiscal policies and extent to which government controlled development (Patel, 1961, p. 323).

Criticisms

Sector theory has come under heavy criticism since its specific delineation by Colin Clark in 1940. Its major premise is that a rising income creates demand which leads to proportionate increase of employment in secondary and later tertiary industries. The reason for this shift is that tertiary and secondary products have an advantageous income elasticity of demand over secondary and primary products respectively. A number of economists have disputed the inevitability with which employment shifts are said to occur.

Richardson points out that sector theory explains growth in terms of both supply and demand. In terms of supply, growth occurs when secondary and tertiary industries offer high returns to capital and labour because these industries' productivity is higher than that of primary industries. Consequently, capital and labour are attracted to secondary and tertiary industries. In terms of demand, growth occurs when rising incomes result in a greater demand for secondary and tertiary commodities because these have higher income elasticities of demand than primary commodities. Capital and labour are consequently shifted into secondary and tertiary sectors. Richardson disagrees with the necessary correlations cited above. He argues that an increase in per capita wages need not coincide with the shifting of workers into secondary and tertiary sectors, nor is there any reason why high income elasticity of demand should accompany increasing production (Richardson, 1969, p. 340).

Bauer and Yamey suggest a number of examples that support Richardson. They point out that growth can be accompanied by a diminution of labour force size in the tertiary sector. A primitive region may employ a large proportion of its labour force in transportation and trade because such facilities are poorly developed. As the region grows, capital may replace labour in tertiary industries as part of the process of automation and technological development. On the demand side, Bauer and Yamey suggest that higher income may result in increased purchasing of primary and secondary rather than tertiary commodities. Examples of primary commodities for which this is true include fur coats, oysters, caviar, lobster, pheasants and orchids. Examples of secondary products include cars, jewelry, works of art and textiles. These indicate that luxury items may come from primary and secondary industries and that certain tertiary industries such as transportation and commerce are necessities with lower elasticities of demand (Bauer and Yamey, 1951, pp. 748-750).

One of the broader criticisms of sector theory is that it over-emphasizes industrialization (North, 1955, p. 245). We would expect this argument from an architect of export base theory. North also makes the general criticism that the theory does not provide us with insights into the causes of growth (1955, p. 245). This point is supported by a number of critics who state that it ignores external stimuli to growth (Richardson, 1969, p. 340; Perloff et al., 1960, p. 60). Thomas agrees, pointing out that the theory begs the question of how to gain the initial increase in income necessary to generate growth (1964, p. 424).

Another set of criticisms deals with the inadequacies of data classification used to test the hypothesis. In early stages of growth, a region's labour force is often classified in agriculture when in fact members spend as much time transporting their produce to market and selling it as they do producing it. Inadequate classification of primitive labour forces creates a false impression that the labour force shifts into secondary and tertiary sectors as the region develops (Bauer and Yamey, 1951, p. 742).

Inadequate methods of classification has also been a criticism of the Hoffman hypothesis. The problem of distinguishing between consumer and capital goods is well illustrated by Hoselitz who states that a towel is a consumer good when a householder purchases it but a capital good when a hotel purchases it. He also argues that use of commodities varies from region to region, making them consumer goods in one region but capital goods in another (Hoselitz, 1961, p. 539). Hoffman's hypothesis has escaped serious criticism except for this minor point. Patel argues that the problem is not serious because capital goods can be defined as all goods used primarily by manufacturers while consumer goods are all those bought by the public primarily for consumption in the home (1961, p. 321).

The Hoffman hypothesis comes off much better than the Clark-Fisher hypothesis. The latter is criticized because it is deterministic and does not explain the causes of growth. It is more a description of the historical process of growth experienced in Europe. However, it has provided us with a number of crucial factors of development. These include the importance of elasticity of demand in an economy where per capita income is rising, industrialization, lead industries, increased division of labour and diversification of industries (Thomas, 1963, pp. 66-67).

Growth Pole Theory

The theory of growth poles places greater emphasis on space than the other two. It is a theory of unequal growth within a geographic space. Thus, growth in one part of a region, called a growth pole or growth centre, determines the extent of growth in another part of that region or in another region. Growth is a function of technical innovations and linkages initiated by a propulsive industry located at the growth centre. Propulsive industries are those with numerous intensive linkages which are capable of expansion independent of the population size in the region. Technical linkages occur when industry A uses commodities produced by industry B. When it becomes economically feasible to do so, industry A may select a site close to industry B, resulting in agglomeration.

A central concept in growth pole theory is the growth pole. It was originally conceived as a purely economic concept by Perroux. He intended it to be a rapidly growing, technologically advanced industry. Its growth was supposed to generate growth in other industries through forward and backward linkages and to provide an environment of growth which would stimulate innovations in all linked industries (Hilhurst, 1972, p. 1). Darwent states that it usually is a heavy, capitalized industry consisting of large firms which benefit from economies of scale (1969, p. 9). Economies of scale result in reduction of costs and maximization of profits due to mass production or reduced input costs.

A growth pole, which is an abstract, economic concept, is found in a growth centre, which is an empirical, geographic term. A growth centre is a point of agglomeration in which one or more growth poles are located. There is a tendency for the growth centre to have strong linkages with the national economy. It may be the centre of a labour market or a major retail trade area. It usually contains high level tertiary functions, carries on a large volume of wholesale trade and is part of a well developed communications system (Darwent, 1969, p. 9).

A number of authors have confounded the two concepts. Nichols defines a growth pole as:

> an urban center of economic activity which can achieve self-sustaining growth to the point that growth is diffused outward into the pole region and eventually beyond into the less developed regions of the nation (1969, p. 193).

In addition to the semantic confusion, Friedmann has expanded on the original meanings by introducing the concept of a 'social development pole' in underdeveloped countries. These poles are centres of concentrated population settlement in which the values and attitudes of the inhabitants are predisposed towards acceptance of change. Relatively higher levels of education prevail and many employment opportunities exist in secondary activities. Social development poles tend to have a fluid, open social structure with leadership elites actively seeking change (Friedmann, 1968, p. 368).

A number of finer distinctions remain after differentiating between growth poles and growth centres. A task force set up by the Canadian Department of Regional Economic Expansion to consider Sept-Iles and Trois Rivières as potential growth poles in Quebec discriminated between poles de développement and poles de croissance. Poles de développement are strong generators of spread effects, for example, Montreal. Spread effects are forces of growth occurring outside the growth pole. Poles de croissance are strong reactors to the development generated in Montreal (Higgins, 1972, p. 226). Higgins discriminates between two types of growth centres. Moore distinguishes between three types of growth poles: (1) the principal pole represented by dominant firms, for example, steel and chemicals; (2) derived poles of utilizers technically linked to the dominant pole, for example, chemicals and paper; and (3) lateral poles or market oriented poles resultant of the growth in income created by the other poles, for example, paper and food products (Moore, 1972, p. 258).

Propulsive industries play a major role in growth pole theory. The concept of dominance is an important aspect of the propulsive industry as formulated by Perroux. Once again Perroux's concept of domination was economic rather than geographic. Domination occurs when "a firm controls an abstract space, the market for a product of a service or a group of products and services" (Hansen, 1967, p. 714). Two industries, I and J, exemplify dominance when the flow of goods and services from firm J to firm I is a greater proportion of J's output than is the flow from I to J of I's output. I is said to be the dominant firm.

The definition of a propulsive industry extends beyond the concept of dominance. Its essential features include the properties of being independent of the population size of the surrounding region thereby releasing products beyond its own territorial limits, and of increasing the circulation of money within its own region. It is supposed to cause "an increase in the size of other firms through capital investment and the introduction of technical innovation. It modifies the economic structure of a given concentration [and] the organization of enterprises" (Vanneste, 1971,

p. 24). Hansen states that the literature emphasizes three properties of propulsive industries: they must be large, rapidly expanding and have a large number of intense interrelationships with other sectors of the economy (1971, p. 222). Moore states that the four major actions of the propulsive unit are: (1) price lowering and anticipation of expansion; (2) leadership and dominance; (3) high propensities to consume, save, invest, innovate, and work; and (4) growth, development, and progress (1972, p. 258).

Growth centres can be conceptualized as agglomerations from which forces of attraction and repulsion emanate. These forces have polarizing effects. They attract investments, industries, innovations, and labour. They repulse stagnant industries with outmoded technical processes and obsolescent products. They attract capital and encourage selective migration of the young, skilled and educated (Darwent, 1969, p. 15). They repulse the unskilled, uneducated and the elderly.

The process of polarization can be applied to many fields. In economic literature, three kinds of polarization phenomena can be distinguished: (1) technical polarization, (2) income polarization and (3) innovation polarization.

Technical polarization occurs when industries providing raw materials to the propulsive industry (backward linkages) and those which consume the product of the propulsive industry (forward linkages) find it most profitable to locate in proximity to the location of the propulsive industry. An important question is which activities exercise the greatest technical polarization. Vanneste suggests that four major factors are:

(1) growth rate of the industry;
(2) the number of phases in the production process;
(3) the intersectorial degree of linkage of an industrial branch;
(4) the diversity of the regional economic structure (Vanneste, 1971, p. 29).

The growth rate of an industry is certain to affect decisions of whether to locate linked industries nearby. A growing industry requires increased supplies of raw materials and provides larger output, thus guaranteeing profits to those linked industries.

The number of phases in the production process is the extent to which a firm's activities are spread over several enterprises within a particular branch of industry. An increased number of phases results in more technical linkages and thus more opportunities for technical polarization.

To appreciate the third point, we must recall that the products of a firm have three possible destinations: (1) they can be used by their own firm, (2) they can be supplied to another firm, or (3) they can go to the consumer as a finished good. A firm whose products are consumed by other firms and/or consumes the products of other firms will have a polarizing effect on the other firms.

Finally, the essence of the fourth point is that technical links will operate most often in a diversified structure rather than in a monolithic structure. A diversified structure offers more opportunities for polarization and promises greater solidarity and stability.

The polarization of income is the Keynesian income multiplier effect. It is the increase in income caused by the flow of expenditure and its effects as well as by the flow of goods into and out of the growth centre. It is influenced by:

(1) the regional consumption quota;

(2) The rate of outflow of money into other regions by attracting outside workers;

(3) the extent to which the profit of the entrepreneur remains within the economic region;

(4) the extension and diversity of economic activities (Vanneste, 1971, p. 39).

Friedmann has emphasized the innovational aspect of polarization, pointing out that channels of communication can have a powerful effect on regions of growth (Friedmann, 1968, p. 365). He maintains that polarization comes about through the concentration of innovations in a "core region" which are diffused to a "peripheral region" resulting in regional development. Core regions reinforce their dominance by means of a feed-back system whereby:

(1) resources are transferred from periphery to core;
(2) increased interaction and innovation occur in the core;
(3) social and institutional changes occur in the core which favour innovation;
(4) innovations tend to induce more innovations;
(5) there is an effect of increased scale and agglomeration in the core region (Hansen, 1971, pp. 27-28).

Lasuen endorses the innovation linkage as the key process in growth of an underdeveloped region. He argues that development is:

> the process by which ... firms adopt successive sets of innovations, the time lapses between which become shorter, and the scale of operations in any one line, larger. From that perspective, the relevant development subprocesses are those relative: (1) to the generation of new innovations [sic], (2) to the diffusion of knowledge of new innovations [sic] and (3) to the spread of adoptions (Lasuen, 1969, p. 144).

The essential consequence of the propulsive industry is that it contributes to agglomeration, which plays an important role in the development of a growth centre. Agglomeration is defined on page twenty-eight as the massing together of firms for reasons of linkage effects and external economies, both of which have been described on pages twenty-eight and twenty-nine. An alternative way of conceptualizing external economies was first described by Scitovsky and has been endorsed by growth pole theorists. External economies are dichotomized into pecuniary and technological externalities. Pecuniary external economies occur when expansion in industry A gives rise to profits in industry B. This may be induced when (1) A produces goods which are inputs to B; (2) A produces goods which are complementary to goods produced by B; (3) A produces goods which can be substituted for goods which are inputs to B; and (4) higher incomes result in larger consumption of outputs of E (Darwent, 1969, p. 7).

Technological external economies are less common than pecuniary external economies. They are the physical influences that are not accounted for through prices. They occur when the output of a firm depends on output and factor utilization of other firms. Examples include the adverse effect of pollution caused by one producer upon another, and the presence of a skilled labour pool because of the location of many firms in one centre (Ebel, 1973, p. 27).

Research

Research of growth pole theory has been primarily descriptive and planning-oriented (Lasuen, 1969, p. 140). There has been little emphasis on either conceptualizing the formal aspects of polarized growth in a spatial setting or determining where growth poles occur (Casetti *et al*., 1971,

p. 377). Only two of the studies to be described in the following paragraphs make an attempt to test hypotheses derived from growth pole theory. All studies that are planning-oriented have been omitted from this section because this aspect of growth pole theory is not relevant to the thesis.

Moore analyzed business philosophies and practices of several entrepreneurs and corporations in the Puget Sound region of the state of Washington. He found a conscious "growth pole philosophy" practiced by the entrepreneurs (Moore, 1971, p. 218). He also compared the development of the metal products complex with that of the forest products complex using a polarized growth perspective. The former complex concentrated on backward linkages and was quite successful although subject to fluctuating demand for its products. The latter established forward linkages but suffered from limited expansion due to the limited market within the region and due to the nature of the product. Wood is not as malleable, strong or light as metal (Moore, 1971, p. 219).

Lasuen attempted to test a basic idea of growth pole theory which is that polarization is generated by innovations. He surmised that innovations would cause "the urban structure to have become very unstable, with the cities experiencing sudden jerky, explosive bursts of growth and decay" (Lasuen, 1969, p. 142). He found the opposite to prevail in European and United States urban systems, although an underdeveloped country, Venezuela, tended toward less stability.

Hodge tried to identify growth centres in eastern Ontario. He used data from eighty incorporated areas in nineteen counties in the eastern part of Ontario that includes Ottawa and the Montreal-Toronto axis. He applied principal components factor analysis to identify seven underlying community traits from thirty-two variables. The seven factors were physical development (housing amenities), proportions French/English, population age, industrial/commercial development, population size, educational level, and compact development (density of population and intensity of capital investment). He used multiple regression analysis to determine the degree of association between factor scores and four indicators of community growth. The four indicators together with their coefficients of determination were: average family earnings, 1961 (0.84); rate of change in population, 1951-1961 (0.80); rate of change in number of retail and commercial establishments (0.64); and rate of change in assessed valuation (0.30) (Hodge, 1966).

Stafford analyzed data from seventy-two urban areas plus the metropolitan areas of Edmonton and Calgary in an effort to identify growth centres in the Alberta urban system.

He used a common factor model to extract three factors from eighteen census variables. The three factors were: business (proportion of the labour force in trade industries and managerial occupations); age-sex structure (child-woman ratio and proportion under fifteen years); and socioeconomic class (education and income). He obtained factor scores for each urban area which indicated the strength of each urban area on each factor. He found no areas that were strong on all three factors (Stafford, 1975).

Casetti, King and Odland used non-agricultural employment data for thirteen cities in the western United States to evaluate the polarization effects of Los Angeles. Using techniques of regression analysis they rejected the hypothesis that raw increases in employment are polarized on Los Angeles but found proportionate increases of employment to be greater as distance from Los Angeles decreased (Casetti et al., 1971).

Nichols applied two empirical tests of hypotheses from growth pole theory to data from the state of Georgia. He regressed increases in per capita retail sales against increases in median income. The regression residuals were intended to indicate which areas received more than their due portion of additional retail sales. These would then be used to test the hypothesis that income multipliers gravitate towards growth centres. However the regression provided a correlation coefficient of -0.0394, which was far from significant, making any study of the residuals very dubious (Nichols, 1969, p. 198). Nichols attributed the low correlation to population movements and changes in American retailing patterns.

The second empirical test was much simpler. Nichols looked at the distribution of increases in median income between 1950 and 1960. Increases in median income were regressed against increases in population because these were thought to be related. The residuals from the regression were then plotted to see which regions had incomes increasing faster than employment. The results appeared to indicate that Atlanta had a polarizing effect on disproportionate increases in median income.

Casetti and Semple did a stepwise regression of population change on distance from arbitrary origins to establish geographic locations of growth poles in southern Ontario. Optimal origins were selected using percentage change in growth, 1951-1961, for all incorporated towns with populations of between one and five thousand in 1951. Growth rates tended to decline with distance from three optimal origins clustered in the shape of a horseshoe around the western end of Lake Ontario (Casetti and Semple, 1968).

Moseley tests the 'trickle down effect' or 'spread effect' of growth poles by analyzing the impact of a growth centre, Rennes, on its periphery, Brittany. He selects a number of variables which he feels to be measures of the level of prosperity of 352 communes in Brittany. These encompass aspects of demographic expansion, gains in the agricultural sector, domestic amenities as measures of the standard of living, and indicators of local economic activity. The variables were subjected to principal components analysis. The first component, accounting for nearly one-third of total variation, loads highly on domestic amenities and measures of immigration. This component is regressed on an independent variable, log X, where X is the distance from Rennes. The coefficient of determination is 0.12 when based on all 352 data points, and is 0.55 when based on the 62 communes within 20 kilometers of Rennes. This suggests that spread effects are restricted to areas within commuting distance of the growth centre (Moseley, 1973a).

Moseley also analyses the spread effects of Haverhill and Thetford on the surrounding area in East Anglia using data collected from several surveys. He finds considerable centripetal migration both to the growth centres and their commuting hinterlands, largely generated by job opportunities within the two centres. Extra industrial activity is generated by increased local consumption. Moseley concludes that spread effects are restricted to the growth centres and their immediate environs (Moseley, 1973b).

Criticisms

The most devastating criticism is that a single growth pole theory does not exist. Each writer has interpreted the writings of Perroux according to his own biases. They have done this because Perroux was not precise in describing his concepts nor did he specify the relationships between the concepts.

A primary problem is the fact that Perroux used economic space instead of geographic space as the framework of his theory. Application of his theory to regional development required adaptation of the original theory to a geographic framework, with the result that variations in the application were inevitable. Growth poles have consequently been interpreted as lead sectors, attraction poles, development blocks and industrial complexes.

Propulsive industries have been confused with key industries and with primary units. Although no writers agree on the essential properties of a propulsive industry, consensus is that it can be identified as being a large industry, rapidly growing and having a quantity and intensity of interrelations with other sectors (Hansen, 1971, p. 22). Perroux emphasized the size of the firm but Boudeville disagrees with him, pointing to Denmark as an example of growth without the aid of a large firm, and to Lorraine where a large steel industry failed to induce growth in the surrounding region.

We do not have a clear definition of a growth pole. Besides the confusion of whether a growth pole is an industry, firm, or place, we are confronted with the problem of determining whether a growth pole is one with past, present, or future growth, and whether it is to be identified in terms of actual or potential growth (Darwent, 1969, p. 5). Hansen argues that even if a propulsive industry is a cause of polarization, it fails to give a satisfactory explanation of agglomeration because we do not know why the industry in question was attracted there in the first place (Hansen, 1971, p. 22). A propulsive industry is one that encourages growth in related industries but there may be many industries fulfilling this requirement but which do not generate growth centres within their regions.

After the growth pole or growth centre has achieved a degree of maturation, its polarization effects are supposed to give way to spread effects wherein growth is generated in the periphery of the growth centre. Nichols questions whether this will happen. He suggests that investors will always prefer to invest in the growth centre because of the advantages of external economies and other socio-psychological benefits. Demand for goods in the growth centre will be greater than in the periphery because the hinterland produces primary products. In conditions of rising per capita income, areas producing secondary and tertiary products have an advantage over other areas because of the higher elasticity of demand for their products. The growth centre will consequently continue to flourish relative to its periphery. Another hindrance to spread effects is the paucity of educated, skilled younger workers who have been attracted to the periphery. The age structure of the periphery is unlikely to attract investment nor will it provide high marginal returns to any investment that is made in the hinterland. These arguments suggest that the growth centre will continue to drain its periphery, minimizing any beneficial spread effects that may obtain (Nichols, 1969, p. 195).

The theory does not explain and cannot predict because it does not identify and quantify specific factors and properties of economic growth (Darwent, 1969, p. 20). It is a loosely-knit set of relations with no precisely related equations linking sectoral and spatial development (Darwent, 1969, p. 25). It neglects crucial factors of economic growth such as the location and spatial distribution of economic activity that might optimize growth and it makes the serious error of treating the growth pole as a closed system (Hurst, 1972, p. 331).

Although Darwent is extremely critical of growth pole theory as a theory he believes that it has much potential.

It promises to provide a dynamic dimension to our present static models of regional structure, as well as injecting a spatial dimension to development theory. In addition, growth pole theory provides a framework for dealing with planning problems although it has not yet given us any answers to these problems (Darwent, 1969, p. 21).

Comparisons

This section deals briefly with the origins of the three theories. It compares and contrasts their frameworks and crucial features.

Sector theory is based on the empirical findings of Clark and Hoffman in the 1930's. However, it stems from a tradition of viewing change as an evolutionary, predetermined process. It was anticipated by Petty, Turgot, Condorcet, Saint-Simon and Comte (Spengler, 1949, p. 86). These writers provide it with a much longer tradition than export base theory, the germ of which was first conceived in 1928 by Robert Haig in his study of the New York region (Thompson, 1965, p. 446). However, export base theory was not developed clearly until a series of articles by North and Anderson appeared in the 1950's. The assumptions in export base theory derive from Keynesian income theory which was developed during the depression of the 1930's.

Growth pole theory is the junior of the three theories.

It was first conceived by Perroux in 1955. Its roots lie in Schumpeter's ideas which identify innovations as the prime mover of a capitalist economy. This theory is the one presently in vogue but it is the least developed of the three and perhaps should not be labelled as a theory.

The processes of growth which the theories describe are not mutually exclusive, but they differ in their frames of reference and in identification of crucial variables. The essential factors of growth according to sector theory is differential income elasticities of demand among primary, secondary and tertiary goods. The crucial process is industrialization which is deemed necessary if the region is to develop. The crucial factor in export base theory is external demand for basic goods in the region. This results in growth of basic industries which attract nonbasic industries. The development of basic and nonbasic industries are components of growth which eventually leads to industrialization.

Growth pole theory diverges from traditional economic approaches to regional growth by introducing sociological concepts such as 'environment of growth' and 'social development poles'. It also attempts to integrate geographic space into the theory and it endeavours to capture the dynamics involved in the growth process. The essential factor in growth is the propulsive industry which encourages innovation and a spirit of growth in the growth centre. It attracts other industries to the centre, achieving agglomeration and external economies.

The three theories differ significantly in their frames of reference. They agree on three points: they assume factor mobility in a capitalist economic system, and they assume constant improvement in transportation and communication throughout the growth process. Export base theory and sector theory require low population densities during the early stages of development. Growth pole theory does not make this specification although an important aspect of development in the growth centre is agglomeration which involves an increase in population. However this population may come from other parts of the region rather than from outside the region. The possibility of shifts in activity within the region marks growth pole theory as distinct from the other two.

The prime movers of growth vary according to the three theories. Sector theory requires a rising income and differential income elasticities of demand among primary, secondary and tertiary industries. Export base theory requires demand for natural resources while growth pole theory requires the presence of a propulsive industry.

The three theories differ in their specification of the initial endowment of the region. Sector theory assumes that the region has no valuable natural endowment. It is an agricultural area engaged in subsistence activities. Export base theory requires that the region contain a natural resource which has not been exploited and which is in demand by other regions. Growth pole theory requires that the region be endowed by a propulsive industry. This would involve any of the factors of production - land, labour, capital, or entrepreneurship - which allow the region to export commodities at a considerable profit. However, the properties of a propulsive industry make it unlikely to be from the primary sector. Growth pole theory begins with a much broader framework than the others. It is applicable to regions at any stage of maturity while the other two theories restrict initial development to regions in a primitive economic state.

Sector theory has the unique requirement that income must be rising in order that differential elasticity of demand cause an increase in production of secondary and tertiary commodities. This condition significantly weakens sector theory because it can be interpreted as explaining economic growth in terms of economic growth. It states that in order to have rising income, we must begin with rising income. The other two theories do not make this assumption, although they make similar assumptions on a smaller scale. Export base theory assumes that demand for an export is increasing the income of a specific base industry while growth pole theory assumes that a propulsive industry is growing at a disproportionate rate because of demand for its produce. However in the latter two cases demand may be a function of improved technology or of advantageous natural endowment. This makes rising income an intervening variable while sector theory treats it as a given.

Export base theory has often been criticised because its prime mover, demand, is an exogenous variable. However, sector theory could be criticised in a similar way. The proportionate increase in production of secondary and tertiary industries depends upon rising income which is also an exogenous variable in the framework of sector theory. A similar weakness in growth pole theory is the unanswered question why a propulsive industry has a particular location and why it has propulsive properties while others do not.

External demand plays a pervasive role in export base theory. It is essential to every stage of regional development. The exogenous variables of the other two theories are not so ubiquitous. Once income begins rising in sector theory we can understand why it continues to do so, given the assumption of unequal income elasticities of demand and the transfer of production factors from primary to secondary and tertiary industries. Similarly, we may not understand why a propulsive industry was located in a region, but once it is there we can understand how it contributes to regional development.

The process of growth differs according to the three theories. As described by sector theory, it involves industrialization wherein the primary industries are replaced by secondary industries which are later replaced by tertiary industries as the major sector in the economy. Development in the export base framework involves growth and diversification of base industries. These industries may be located in any one of the sectors. Industrialization is an essential part of development of a mature economic region but it is considered a function of external demand. Industrialization occurs only as part of the maturation and diversification of the export base. It is a dependent rather than an autonomous aspect of the final stages of development.

Growth pole theory is not concerned with industrialization other than to emphasize innovations in the industrial process. Innovations can also occur in the social and administrative organization of the firm, as well as in products, raw materials and transport and storage facilities. Innovations strengthen the attracting power of the region causing agglomeration and external economy effects to create a growth centre which eventually leads to spread effects in other parts of the region.

A number of aspects of growth are recognized by all three theories but they differ in the weight they put on each component of growth. Linkages are considered essential to agglomeration in growth pole theory but they also play an important role in the diversification of an export base. Export base theory is concerned with the attraction of forward linked industries because the resident industries are resource oriented, offering little opportunity for backward linkages. Linkages play a role in sector theory as well since it is the nature of secondary and tertiary industries to have more linkages than primary industries (Darwent, 1969, p. 24).

A serious contradiction between sector theory and export base theory is the concept of income elasticity of demand. Export base theory states that commodities of an export base must have high income elasticities of demand. Since the base begins as a primary industry, this requires that commodities of a primary industry have high elasticity of demand. But the crux of sector theory is that the region must develop its secondary and tertiary industries because products of primary industries do not have high elasticities of demand. Some reconciliation can be attained if we recall that sector theory assumes no valuable natural endowment in the region undergoing development. The region is assumed to be an agricultural area which export base theorists recognize as being undesireable as a base because such products have low income elasticities of demand. Growth pole theory endorses elasticity of demand as one of the essential properties of the products of the propulsive industry but it is not emphasized in the literature.

The multiplier effect is crucial in export base theory. It is used to explain the importance of basic industries over nonbasic industries. It is also crucial to growth pole theory where it is conceptualized as polarization of income. The concept seems to be neglected in sector theory.

Summary.

An attempt is made in this chapter to describe each of the three theories and to summarize the literature which deals with research and criticisms of each theory. The chapter ends with comparisons of the crucial similarities and differences among the three theories.

Export base theory is treated first. Its main thesis is that a region grows as a function of its exports. It is based on the North American experience as seen by North, Perloff, and Tattersall. It derives from Keynesian economics whereby investment is a prime mover of the economy.

Export base theory assumes a capitalist system operating in a region of low population density and containing a valuable resource endowment. Prerequisites for growth include demand for the resource from outside the region, and advantages of production and transportation of goods compared to other regions with the same resource endowment.

A serious criticism of export base theory is that a principal variable, demand for the resource, is treated as an exogenous variable in the theory. Other criticisms are that nonbasic activities and terms of trade are not given due recognition. Blumenfield argues that the distinction between basic and nonbasic should be made in terms of whether or not the industry must compete with other industries outside the region to survive rather than in terms of whether or not the industry attracts money into the region.

The main idea in sector theory is that a region grows as a function of a change in its industrial structure. The major economic activities shift from primary to secondary and from secondary to tertiary sectors as a result of greater income elasticity of demand attributed to manufacturing and service industry products. This is known as the Clark-Fisher hypothesis. Sector theory assumes a context of rising per capita income which allows greater consumption of secondary and tertiary commodities.

According to sector theorists, industrialization, which is a crucial process in the theory, involves more than just an increase in manufacturing. Within that sector there must be a rising proportion of activities in the production of capital goods vis-à-vis consumer goods. The advantage of increased capital-goods production is increased specialization, more linkages, and a higher input of capital, all of which result in greater output per unit input. Hoffman was the originator of the idea that capital-goods industries play a crucial role in industrialization.

A major criticism is aimed at the causal relationships prescribed by sector theory. Critics argue that rising income does not necessarily shift economic activities to secondary and tertiary sectors, nor does high income elasticity of demand necessarily accompany increasing productivity.

Another set of criticisms deals with inadequacies of industry classification. It is argued that distinctions between primary, secondary, and tertiary categorizations are not reliable indicators of the real economic activities that occur in developed and underdeveloped countries.

The term 'growth pole' was coined by Perroux in 1949 (Darwent, 1969, p. 5) although the roots of the concept lie in the Schumpeterian system which emphasizes entrepreneurial activity as a force in economic growth. Linkages and propulsive industries also play crucial roles in development of a growth centre which is the geographic location containing the growth pole which is a rapidly growing technologically advanced industry. The essential idea in growth pole theory is that an expanding firm attracts other expanding firms to a growth centre where economic growth occurs as a function of the ensuing entrepreneurial activity, external economies, and economies of agglomeration and scale. Eventually this growth which begins as a centripetal activity generates 'spread effects' or 'trickle down effects', causing economic development in the surrounding regions.

Growth pole theory is in its adolescence. Its concepts have not yet been clearly delimited nor have they been molded into an integrated, structural whole. The theory does not explain why growth begins in one centre and not another, nor why a particular industry is a growth pole. The process of 'spread effects' has been seriously questioned by researchers.

Chapter 3

Methodology

The purpose of this chapter is to describe and discuss methods used to test hypotheses from the three theories described in the preceding chapter. The plan of the chapter is to select the best operational definitions of key variables from each theory and to describe the method of multiple regression analysis by which the variables can be tested to determine which account for most of the variation in average urban income from 1951 to 1961. Operational definitions are selected to closely approximate the concepts, to be easily quantifiable, and to be available in published data. Operational definitions that have already been established will be selected because the task of developing new definitions is beyond the limits of this thesis.

The selection of indicators or operational definitions in this study is based on the research and ideas of earlier studies which investigate the same concepts we address. If more than one indicator has been chosen in different studies in the past, an attempt is made in this study to select the one most appropriate in the context of the research framework and according to the availability of Canadian data.

The link between theory and fact is a tenuous one. The selection of indicators from the real world of data to represent variables from the abstract world of theory is based on what Eddington calls "pious opinion" (1933, p. 255). In physics, mass is measured or inferred by a pointer reading. In sociology, social class is measured by years of completed education. In psychology, inherent intelligence is measured by a 20 minute paper-and-pencil quizz. The appropriateness of these measures as indicators of associated concepts is questionable but custom has deemed them acceptable.

Operational Definitions of the Export Base

The crucial internal variable in export base theory is the economic base which is defined to encompass all economic activities which provide goods, services, or capital to persons or organizations with sources of payment outside the boundaries of the area under analysis (Andrews, 1955b, p. 361). All other economic activities, referred to as nonbasic or service activities, provide goods, services and capital to persons and organizations whose source of payment lies within the boundaries of the area. The base is supposed to be the vehicle of growth in the area in that it attracts investments and income. Service industries support basic industries. They are an essential part of any community but they do not represent the driving force behind growth of that area.

Output of most firms is consumed both internally and externally. The problem then becomes one of determining what portion of a firm's output is exported. If any portion of an industry's output is consumed externally, it is called a basic industry although the major concern is with measurement of the portion exported.

At times there are difficulties in conceptually distinguishing between basic and nonbasic industries, even though they appear to be easily distinguishable. An example of this can occur when two linked industries are located in the same area. If industry A provides goods for industry B which then exports a composite good including the finished product of A (for example, a part of a motor), then the classification of A as a nonbasic industry becomes dubious. This problem would be resolved if the question was directed at the balance of payments and if units were in terms of dollars, in which case only direct sales would be considered.

> Nobody thinks of including the steel industry in the export statistics of American automobiles, because the "value added" by the steel and all other "indirect primary" industries is, by definition, included in the gross value of the automobiles. Dollars, not persons employed, are the correct yardstick (Blumenfield, 1955, p. 124).

A number of tertiary industries are conceptually difficult to classify. For example, a university provides a service to students from outside the urban area, making it a basic industry within that area. Yet the service is consumed within the base area. Students consume many other nonbasic goods while attending university and a large number of graduates obtain employment and residence within the urban area containing the university. The classification of a university as basic becomes somewhat arbitrary (Andrews, 1954b, pp. 262-263). In most base studies, staff of universities are allocated according to ratio of local and nonlocal students. But this is a division according to who benefits, not according to the balance of payments. The real criterion should be the source of the professors' wages (Blumenfield, 1955, p. 125). Does the money come from federal or provincial grants, or does it come from city funds? The latter case makes the university a service industry but the former makes it a basic industry. Other activities that are difficult to classify include tourist facilities, medical and research centres and legislative facilities. Transportation and communication industries are also conceptually difficult to classify. Most interurban transportation is arbitrarily classified as basic even though it performs a service for other industries (Blumenfield, 1955, p. 124).

Commuters tend to confound the problem of classifying basic and nonbasic industries. Their purchases may be basic or nonbasic depending upon the classification of their location according to place of work within the base area or place of residence outside the base area. A similar problem is invoked by the phenomena of chain stores and branch offices whose profits are drained off to other localities.

The difficulties in differentiating conceptually between basic and nonbasic activities are not as great as those inherent in discriminating between the two in the empirical world. Few firms export all their produce but many export a portion thereof. A major problem becomes one of determining what portion of a firm's output is exported.

Delineation of the boundary of the base area plays an important role in determining whether an industry is basic or nonbasic because classification depends on whether output passes over the boundary. The boundary of the base area should be determined by the geographic distribution of capital, labour and services associated with the urban area (Andrews, 1954c, p. 311). Unfortunately such distributions are not coterminous. The distribution of telephone services seldom coincides with that of electric power or with the residential area of workers. The divergence of these distributions leaves the economic base analyst with the choice of selecting one of the two distributions which best represent the concept of base area - the distribution of firms or the labour market area (Andrews, 1954c, p. 312). Andrews implies that the Metropolitan Area comes close to approximating the "theoretical ideal concept" of an economic base area (1954c, p. 316).

The selection of an adequate unit of measurement is a crucial problem to be added to those of delimiting the area and conceptually distinguishing between the two types of industries. Andrews (1954a, p. 53) suggests six alternate units of measure:

(1) employment;
(2) payroll;
(3) value added;
(4) value of production;
(5) physical production;
(6) dollar income and expenditure accounts.

Employment is the most widely used measure, primarily because these data are more readily available than any of the others. Use of this measure attenuates the result of an economic base study for a number of reasons. Variation in output per worker occurs across industries, occupations and time, contrasting with the assumption that number of workers reflects productivity or profits.

The disadvantage of using payroll to measure the basic-nonbasic ratio is that it equates one job of $10,000 income with two jobs of $5,000 each. The ratio of productivity of jobs is not necessarily equivalent to the ratio of their wages. This measure also confronts the problem that workers are paid in ways not involving wages and it omits owners of businesses and unpaid workers. At first glance, the payroll measure has the advantage of reflecting consumption of the worker, which is a valuable guide to the worker's contributions to the economy. But problems arise when cost of living and consumption patterns vary from one urban area to another. In this regard, there is also the circumstance of differential expenditures among different wage brackets. Those persons in lower wage brackets spend a higher portion of their income on consumption than those in higher brackets.

The use of value added and value of production pose difficulties in collection of data. Beyond this dificulty is the distortion caused by complex price movements as well as the task of evaluating intangible outputs of such industries as education and medicine. The latter problem is also a major drawback when physical production is used as the measure.

Dollar income and expenditure accounts embraces all monetary transactions of the community. It represents the most complete and satisfactory statistical view of the urban economy but such an analysis is beyond the capabilities of most base studies. Comparing the strengths and weaknesses of the six measures reveals employment to be the most appropriate single measure considering the relative costs of each (Andrews, 1954a, p. 59).

Analysts who attempt to measure economic bases must be cognizant of the fact that the basic-nonbasic ratio is in a constant flux. Variations may be caused by national or regional economic cycles or by shifts in industrial location patterns reflecting changes in markets, raw materials or transport rates. Three other cuases of change in the ratio are technological and institutional innovations and secular change (Andrews, 1955a, pp. 146-149).

We are not concerned as much with discriminating among causes of fluctuations of the basic-nonbasic ratio as we are with determining the relationship between this ratio and development of the urban economy. A premise in export base theory is that a large economic base presupposes economic development. A simple, obvious method of testing such an hypothesis is to determine to what degree the 1951 basic-nonbasic ratio contributed to growth during the ensuing decade. However, growth from 1951 to 1961 is also a function of concurrent economic activities. Thus, relocation of an army base or a series of summers of bad weather in regions of agriculture or tourism will affect the basic-nonbasic ratio and consequently the economic development of the region. In order to allow for this possibility we must examine the possible interrelationship of regional growth and change in the ratio from 1951 to 1961.

A number of methods have been developed to measure the urban economic base. Two of the most common measures are the minimum requirements and the location quotient. The two differ only with respect to the method of determining how much employment is necessary in an industry to supply the local market. The location quotient takes a national average or some variation thereof while the minimum requirements method selects the city with the lowest portion of its labour force in that industry as the amount necessary for local sustenance. If a community produces more of a good than the derived benchmark, that good is deemed an export item. The assumption is made that if more goods are being produced, more workers must be employed in the industry that produces the good. The proportion of the goods being exported is equivalent to the proportion of workers exceeding the national average. If an industry shows a proportion less than the derived benchmark, it is assumed that its products are imported into the community.

The basic formula for the location quotient is

$$\frac{\dfrac{e_i}{e_t}}{\dfrac{E_i}{E_t}}$$

where:

e_i = local industry employment;

e_t = local total employment;

E_i = national industry employment;

E_t = national total employment.

Thus when applied to a single industry, the measure is a simple ratio of the industry's share of local employment relative to the industry's share of national employment. The minimum requirements method substitutes M_i, M_t for E_i, E_t where M_i and M_t are local industry employment and local total employment in the city with the lowest proportion of workers in that industry, or some variation thereof. The benchmark may be placed at the fifth percentile rather than at the first to allow for exceptional cases.

The minimum requirements method was developed by Ullman and Dacey because the location quotient was said to underestimate exports and consequently to overestimate the multiplier which is often calculated from the quotient (Ullman, *et al*., 1971, p. 16). The overestimate occurs because measurement from a national average results in the inclusion of exports to the urban hinterland in the nonbasic portion of production. Almost every city is required to produce a certain amount of goods for its service area beyond the city limits. The location quotient estimates fewer exports than there really are because most cities export goods. The average city thus exports goods although the location quotient equates the average with the amount necessary to meet internal needs. McCalden attempts to circumvent this problem by determining the amount of produce that is consumed by the service area but this is feasible only when the investigator is capable of surveying commodity flows between the hinterland and the city (1973).

Concern for the multiplier is warranted when it is derived from the basic-nonbasic ratio but such is not the case in the present study. What is needed is an indicator of the size of the economic base. If the location quotient contains a bias, it does so for all cities, thus retaining viable comparability.

There are a number of assumptions in the location quotient coefficient and minimum requirements method which are erroneous and which weaken any inferences which can be made from them. The first assumption is one of uniform demand among all cities. Residents of one city do not consume the same quantities of goods as residents of other cities. Winter clothing and heating fuels are utilized more by residents of cities in colder lattitudes. Income levels vary from city to city, dictating different propensities to consume different goods (Isard, 1960, p. 125). Tiebout suggests that to some extent these variations can be accounted for if data on consumer expenditures are available for each city. However these are broken down by the products people buy rather than by the industry producing the product, making it tedious to inculcate such data into an economic base study (Tiebout, 1962, p. 48).

A second assumption open to criticism is that production practices are constant from city to city. This is not always true, as cited by Isard who points out that in the production of steel ingots, the United States Pacific Northwest uses more scrap per ton steel than the Great Lakes region. Thus the location quotient coefficient and the minimum requirements method for the steel scrap industry, which is higher in the Pacific Northwest, does not reflect real production differences between the two regions (Isard, 1960, p. 126).

Labour productivity is a type of production practice that will also vary from city to city. Criticism of this assumption is more significant when we realize that the location quotient coefficient and the minimum requirements method measure the industrial labour force to estimate industrial productivity. Tiebout recommends that when available, data on value added, which is an indicator of productivity, should be used to weight local employment for input into the location quotient coefficient (Tiebout, 1962, p. 48). The same should be done when the minimum requirements method is used.

A third criticism deals with 'industrial mixes'. Different regions have different combinations of industries. Thus industries utilize each other's products so that some products which are defined as exports by the location quotient coefficient and minimum requirements method are really consumed within the region. For example, a region with a large number of packing plants is also likely to have a large number of tin can manufacturers. The cans are sold locally but the location quotient coefficient and minimum requirements method indicate that they are exports (Tiebout, 1962, pp. 48-49). However Tiebout points out that this may be a strength of the two measures rather than a weakness in that they measure indirect exports (1962, p. 49). Presumably the exports of industries using tin cans increase because of the availability of these local commodities. However Blumenfield provides an exceptional example of an exporting branch plant which imports most of its raw materials. If part of its produce is sold locally, displacing products of a firm which uses local raw materials, the net effect is for the balance of payments of the region to move in the opposite direction to that indicated by the location quotient or minimum requirements method of the particular industry (Blumenfield, 1955, p. 123).

A final inadequacy of the location quotient coefficient and minimum requirements method is the fact that they are functions of the breadth of the industry classification being used. A broad classification may combine an exporting with an importing industry, resulting in a location quotient coefficient of about one. Thus the coefficient indicates no exports in the broad industry category. But when data are available for each of the sub-industries, exports may be revealed. This problem has been given the label of 'product mix'. It reveals that the narrowest possible industry categories should be utilized in calculating the location quotient coefficient or the minimum requirements method.

The minimum requirements method is less desirable of the two in this study because the limited size of some of the cities increases the possibility for capricious variations to creep into the data. The benchmark of the minimum requirements method is placed at one end of the distribution, making it more sensitive to such variations than the location quotient coefficient in which the benchmark occurs at the mean of the distribution. The minimum requirements method is best used when metropolitan areas are being compared. The labour force size is then large enough so that unusual divergence from the national average is less likely.

The location quotient is used in this study as an indicator of the size of the economic base. The formula used here is also used by Crowley, Grossner, and Maxwell, and is based on a formula developed by Ullman and Dacey (Crowley, 1973, p. 92; Grossner, 1970, p. 40; Maxwell, 1965, p. 87; Ullman and Dacey, 1960, p. 89). The formula is

$$S = \sum_{i=1}^{n} \left[\frac{(m_i - M_i)^2}{M_i} \right] \Bigg/ \left[\frac{(\sum_{i=1}^{n} m_i - \sum_{i=1}^{n} M_i)^2}{\sum_{i=1}^{n} M_i} \right] \qquad m_i > M_i$$

where:

$$m_i = \frac{X_{ij}}{\sum_{i=1}^{n} X_{ij}}$$

$$M_i = \frac{\sum_{j=1}^{r} X_{ij}}{\sum_{i=1}^{n} \sum_{j=1}^{r} X_{ij}}$$

x_{ij} = employment in industry i in city j

n = number of industries (56),

r = number of urban areas (46).

Thus m_i is the proportion of city j's labour force engaged in industry i while M_i is the proportion of the aggregate urban labour force engaged in industry i. Note that the only industries that figure in this calculation of S are those for which the city proportions are greater than the aggregate urban proportions. In the formula, the difference between the two proportions for each relevant industry is squared, divided by the aggregate urban proportion, and summed for the city. Squaring the difference in proportions increases the effect of industries with larger location quotients. This number is then divided by a value which "weights down large deviations in specialization and accentuates larger numbers of small deviations" (Grossner, 1970, p. 40). In other words, cities with one or two very large industries are weighted down more than cities with several reasonably large industries. This divisor is the ratio of two numbers. The first is obtained by squaring the difference between the sum of the proportions of the city labour force and the sum of the proportions of the aggregate urban labour force in all relevant industries. The second number is the sum of the proportions in the aggregate urban labour force.

Operational Definitions of Industrial Sectors

The following three operational definitions of sector growth are relatively straightforward. The major thesis of sector theory is that a region develops because of a percentage increase of the labour force employed in (1) manufacturing and (2) service industries, and because there is an increase in proportion of persons engaged in (3) capital goods industries within the manufacturing sector.

We shall define these terms. Manufacturing industries are easily identified and are contained within the broad classification entitled 'Manufacturing'. Service industries can be defined as those economic activities which "take the saleable form primarily or exclusively of a personal service rather than a material commodity" (Stigler, 1956, p. 47). In the present study, all industries falling under the broad classification of transportation, trade, finance, insurance and real estate, services, and public administration are classified as service industries.

The identification of capital-goods industries is more difficult. The author of the hypothesis that capital-goods industries must grow proportionate to consumer-goods industries is Walther Hoffmann. He defines consumer goods as

> those finished goods which are ready for immediate consumption and also those semi-finished goods which although often used in

> industry, are largely bought by the public in a finished form primarily for consumption in the home.... Producer goods [capital goods], on the other hand, are those products which are used in a finished or a semi-finished form by industry alone (Hoffman, 1955, p. 73).

Hoffmann suggests that those commodities which are clearly consumer goods are food, drink, tobacco, textiles, clothing, leather goods and furniture. Commodities which are almost entirely capital goods are produced in such industries as chemicals, vehicle building, machinery, tools and scientific instruments, and ferrous and nonferrous metals and hardware. The remaining manufacturing industries are not classifiable as capital-goods or consumer-goods producing industries because their products are purchased by both consumers and other industries. These industries include rubber, building, timber and woodworking industries, paper, paper goods, printing and publishing, water, gas and electricity, and quarrying, cement, glassmaking and pottery (Hoffmann, 1958, pp. 8-9). Taking a cue from Hoffmann, we shall restrict classification of capital-goods industries to those whose produce is purchased on a negligible scale by consumers. These include the broad manufacturing sectors of metal and machinery, transportation equipment, and chemicals.

Operational Definitions of Crucial Variables in Growth Pole Theory

The two variables selected as indicators of the strength of growth pole theory are agglomeration and entrepreneurship. According to Hoover, agglomeration consists of (1) economies of scale, (2) localization economies, and (3) urbanization economies (Isard, 1956, p. 172). Scale economies are those applicable to growth of a single firm. Economies of location are those incurred when firms in a single industry mass together, while urbanization economies are those which occur when firms in all industries locate in a single place.

Population would increase concomitantly with agglomeration. This is superficially the most appropriate operational definition available. However, its efficacy as an indicator of agglomeration is confounded by variations in the age-sex structure of urban populations. Such variations are not too serious but can be reduced if we select labour force size rather than total population as the operational definition of agglomeration. Labour force size reflects the number of jobs available, which makes it a valid indicator of urbanization economies which are an easily measured part of agglomeration. The main disadvantage of labour force size is its sensitivity to annexation. However this influence will be present whatever definition is chosen.

The second variable taken from growth pole theory is entrepreneurship. It may be defined as an act of coordinating and controlling factors of production and instituting major innovations (Hurst, 1972, p. 339). The principal manifestations of entrepreneurial activity involve the setting up of new establishments, expansion of existing ones, and introduction of innovations.

Innovations may be defined as the commercial introduction of new or improved products or processes. This can be contrasted with the definition of inventions which are discrete technical advances comprising new combinations of existing knowledge for practical use in production (Pred, 1966, p. 89). The distinction between the two is that an invention broadens the knowledge of possible technical advancement while an innovation represents an attempt at technological advancement.

Patents have traditionally been used as indicators of entrepreneurial activity, but we should be aware of the tenuous relationship between the two. Patents may be considered indicators of inventions which are associated with innovations which in turn are only one aspect of the entrepreneur's function. Note Pred's statement that "some patents are no more than a legal scrap of paper representing nothing but foolishness and chicane" (1966, p. 108). The incidence of patents may not be a good reflection of inventions because a number of inventions do not get patented. Either the patent may be denied or the individual may wish to avoid the cost of patent procurement. Individuals or corporations may prefer not to patent an invention for fear that revelation of cost-cutting production techniques may eliminate market expansion opportunities (Pred, 1966, p. 109). Many patents never become innovations while many innovations do not stem from patents (Pred, 1966, p. 110).

Further limitations result from the discrepancies in time and location between the granting of a patent and its implementation. The implementation may occur before or after the patent is granted, and it may occur in a city distant from where the invention was made. This latter point considerably weakens the utility of patents as indicators of innovation when differential urban development is the prime consideration, as it is in this case. Patent data provide us with the residence of the person granted the patent, but not the location where it is applied. Identifying the residence of persons who are granted patents is of no use if the application of the invention occurs in a different city.

A final weakness in the use of location of patents to measure location of innovations is the considerable variation in the scale and impact of innovations. All patents must be given equal weighting although some are never used, some prove uneconomical when applied, others are of marginal use, while a few become revolutionary innovations of considerable economic gain to the initiators.

We reject the incidence of patent procurement as an operational definition of entrepreneurial activity for the preceding reasons in spite of its popularity in other studies. The remaining facets of entrepreneurship include expansion of old firms and establishment of new ones. The latter component is the only one for which data are available. The data - number of retail stores, wholesale trade, and service trade locations in 1951 and 1961 - provide us with the net change in number of these establishments over the decade. Manufacturing establishments could not be included because such data are not available for several towns in the study. The Dominion Bureau of Statistics did not publish manufacturing data if a single firm was responsible for more than 75 percent of gross earnings or if two firms were responsible for more than 90 percent of the gross earnings in that urban area.

Unit of Analysis

The purpose of this thesis is to explain differential growth of Canadian cities. The units of analysis are consequently selected Canadian cities which must be gauged at two or more points in time to capture the dynamics of growth. Recent data are desirable because they are more relevant to predictions in the future and because more refined industry data necessary for calculating location quotients are found in later censuses.

The most recent data appear in the 1961 Census. It seems appropriate to use data from 1961 and one or more earlier censuses whose data are comparable with those of 1961. Reclassification of industries by the Dominion Bureau of Statistics was necessary from census to census to account for establishment of new industries and for new technological developments. Thus, only about 55 percent of the 1961 classes of industries are directly comparable with 1951 and only 35 percent with 1941 (Dominion Bureau of Statistics, 1966b, p. ix). Fortunately the Division of Industry Statistics, Statistics Canada, provides convertibility indices for the 1948 and 1960 Standard Industrial Classifications. The 1948 SIC was used to code industries in the 1951 Census while the 1960 SIC was used for coding industry data from the 1961 Census.

The 1941 Census data were not included in this analysis for two reasons. In 1941 Canada was engaged in war with consequent disruptive effects on its industrial structure. This makes comparisons of 1941 with 1951 or 1961 industrial data less viable.

A second difficulty arises from changes in the census questions regarding employment. In 1941, respondents were asked if they had been 'gainfully employed' at any time in the preceding year while in 1951 and 1961 they were asked to specify their type of employment during a particular week. The question of gainful employment attempts to get at the person's usual employment which is difficult to compare with his activity in a selected week of the year as discerned by the latter question (Dominion Bureau of Statistics, 1966a, p. xiv). These discrepancies influenced our decision to exclude 1941 Census data from the analysis.

The 1951 and 1961 data are easily comparable although two minor discrepancies exist. Indians on reserves were classified separately in 1951 but not in 1961. This change does not affect the study because no metropolitan or urban areas contained Indian reserves in 1951. Quebec, Montreal, and Vancouver M.A.'s were enlarged to include parts of Indian reserves in 1961 but the change in classification of Indians on reserves does not affect these data.

The other discrepancy is the change in minimum age in the labour force data from fourteen years in 1951 to fifteen years in 1961. Fourteen year olds play a minor role in industry and can be ignored under the assumption that their exclusion in 1961 did not differentially affect the industrial structures of cities to a significant degree.

The most desirable unit of analysis is the area corresponding to the distribution of residences of the labour force associated with the urban area, but the available data pertain to areas described by political and census boundaries. The discrepancy, where it occurs between administrative boundaries and the actual distribution of the population is of concern because human ecology studies reveal that the urban population does not distribute itself homogeneously but forms heterogeneous clusters each of which is internally homogeneous. The comparison of cities is troubled when dissimilar groups reside outside the political boundaries of cities and when some city boundaries encompass their total population while others do not.

Comparing urban units over time is confounded by annexations. If an area annexed between 1951 and 1961 was uninhabited in 1951, the annexation has no effect on the change in the values of the indicators; however, if the area contained a large population in 1951, the population would distort the values of the indicators within the urban area. This is because the presence of one or more homogeneous groups outside the urban boundary causes the values for the urban area to be determined by only part of the total industrial distribution. If that homogeneous group is a necessary part of the city, its absence will distort the data.

In most cases, data were available only for areas delimited by political boundaries and in many cases the political boundaries were changed between 1951 and 1961. We were forced to make the assumption that values of indicators for the political areas were equivalent to values for the total population. This assumption is not necessary for cities of over 100,000 population for which data were published for the metropolitan area which included the city and all adjacent populations predominantly engaged in non-agricultural employment.

Although metropolitan areas represent ideal units of analysis they pose the problem of not being completely comparable with urban areas which are delimited differently. The attenuation inherent in comparing metropolitan areas with urban areas was considered less than the alternative of using the urban area part of the metropolitan areas and of consequently losing the large suburban populations of the cities over 100,000. Those units which gained metropolitan area status between 1951 and 1961 are treated as urban areas for both census years because data for areas equivalent to the 1961 M.A. boundary could not be obtained for 1951.

The minimum size of the urban units selected for this study was determined by data availability. The Dominion Bureau of Statistics published labour force data for all incorporated areas of 1,000 or more for 1951 and 1961. More detailed data were published for areas of 10,000 or more, data in still greater detail for areas of 30,000 or more and finest detail for metropolitan areas. Location quotients become more valid as industry categories are narrowed because broad categories do not reveal shifts of industry structure within those categories (Anderson, 1968, p. 6). The categories for cities of 30,000 or more were considered adequately narrow to compute valid location quotients but this restricted us to thirty urban areas. We were able to increase this number to forty-six by using unpublished data which Grossner had obtained and thoughtfully included in an appendix of his thesis (1970, pp. 105-151). A list of the fifty-six industry categories is found in Appendix A.

Grossner's data apply to all urban areas with populations in excess of 25,000 in 1961 with the exception of Corner Brook, Newfoundland and Noranda-Rouyn, Quebec for which 1951 data are not available. At that time Corner Brook was unincorporated while the population of Noranda-Rouyn was less than 10,000.

In cases where two or more urban areas are located in proximity with each other, they are treated as one unit of analysis. Appendix B contains a list of the units of analysis and their corresponding urban areas of metropolitan areas.

Operational Definition of Urban Development

A satisfactory measure of urban growth is not available to the urban economist (Richardson, 1971, p. 79). Growth in the national economy is measured in output or income terms but this kind of data is not available at the urban level. In consequence, urban growth is often measured in terms of an increase in a city's population. Such a measure runs the risk of confusing growth and welfare criteria. In large cities there is some evidence that expanding populations bring about a decline in welfare (Richardson, 1971, p. 80). This is because of diseconomies of scale associated with pollution, housing shortages, rising living costs, and excess demand on social overhead capital and welfare services.

Siebert suggests that regional development can be viewed in either of two ways. It can be considered as an increase in income accruing to residents of the area or as an increase in the set of final commodities available in the region (Siebert, 1969, p. 4). This thesis shall pursue the former measure which is much easier of the two to operationalize at the level of urban areas.

The concept of regional income stems from adaptations of the concept of national income (Isard, 1960, p. 82). Kuznets states that national income is "the net product or net return on the economic activity of individuals, business firms, and the social and political institutions that make up a nation" (Isard, 1960, p. 83). It may be viewed as (1) total income earned by owners of productive factors: wages, interest, rent, and profits, or as (2) net value of all goods and services produced in a given year. The latter approach provides a measure of net national product which involves the summing of personal consumption expenditure, net private investment, and government purchases of goods and services (Samuelson, 1955, p. 192).

National accounting systems are based on the net value approach. However, regional accounts are considerably more difficult to derive. At the national level, political, economic, and geographic boundaries often coincide, encouraging aggregation of data at this level, but regions are not so blessed. They may be delimited to coincide with economic or geographic lines but be non-coincidental with political boundaries, making details of input and output more difficult to measure. Transactors tend to differentiate in their records between internal and external trade at the national level but do not do so with respect to intra- and extra-regional trade.

Many transactors, such as many large corporations and provincial and federal governments, are spread over several regions. It would be inaccurate to assign each of them to a single region and difficult to determine what portion of each is internal to each region. A branch plant may contribute to the economic growth of a region but if its head office is elsewhere, the branch plant profits cannot be credited solely to the region's account. The region containing the head office is also a beneficiary and if the corporation is public, stockholders throughout the country can be said to benefit from profits of the branch plant.

Regional accounts are made more tenuous by modern modes of transportation which allow commuters to travel considerable distances to work. If they pass over a regional boundary on their way to work, a record of their wages is misleading when credited either to their home region or to that of their work. Their income should be credited to the area where they make their purchases whether it be home, work, or a third region.

To summarize, the two approaches to regional income are total earned income and net value of all produce. The first of these will be used in this thesis because data on personal and government purchases and private investment, which are the measures that make up net value of all produce, cannot be obtained by urban area.

Measures of total earned income are not available but wages, which consistently make up about two-thirds of all earned income, can be obtained (Samuelson, 1955, p. 186). Such data are available for all urban and metropolitan areas in Canada for both 1951 and 1961 Censuses.

Vida Nichols uses residual change in median income as a measure of regional development. His argument is that wage levels are a function of supply and demand of labour and of the degree of skill required (Nichols, 1969, p. 198). Thus a rise in wages probably means an increase in demand for labour or for more skilled labour. Either case indicates an expanding economy in the region where it occurs.

Canadian census data provide mean rather than median wage for 1961. Calculations of median wage would have had to be based on broad wage categories, so it was decided to retain mean wage from 1961 and to calculate corresponding means rather than medians for 1951.

Operational Definition of Change

Change is a second-order abstraction (Coleman, 1968, p. 429). The first order occurs when we develop concepts of the state of our environment through our sense impressions. The second-order abstraction occurs when we compare two sense impressions and simultaneously consider the length of time between formation of these impressions. This additional abstraction has delayed vigorous conceptualization of change in the social sciences.

Any attempts to deal with change in the social sciences have been met by a number of problems arising from the measurement of such phenomena. The following paragraphs deal with the problems encountered and mention four ways of measuring change. The most satisfying method will be selected for use in this study.

The first problem was noted by Thorndike in 1924 when he pointed out that there was consistently a negative correlation between the initial score on a variable and subsequent change in that variable (Bohrnstedt, 1968, p. 116). Spurious negative correlations may result from two causes. These are regression effects and what Bereiter calls 'over-correction-under-correction' (1963, p. 3).

The concept of a regression effect originated when it was observed that children's height tended to be closer to the mean population height than did those of their parents. Other studies have found this phenomenon to occur on a number of other variables. For example, students writing pre- and post-tests have been found to demonstrate this behaviour. Scores of those who excelled and those who failed tend to regress toward the mean on the later exam. Coleman suggests three possible reasons why regression effects occur. The first he labels 'negative feedback' which is the equilibrating process in systems theory whereby a value, upon finding itself far from the mean, will tend to return. The other two reasons listed by Coleman are assumptions implicit in systems theory. The first is that variables in a system are in aggregate equilibrium and consequently individual values must return toward the mean if they deviate enough to destroy system equilibrium. The second is that each individual is characterized by the same values of parameters which govern the system (Coleman, 1968, pp. 438-439). Thus every variable is influenced by the same mean and variance as the system as a whole.

The regression effect is not universal. One obvious exception is the comparison of population size over time. Here we find high positive correlation between initial population and subsequent change but the correlation derives from the use of a linear model of growth when an exponential model should be used.

Regression effects refer to changes in real values. Over-correction-under-correction refers to the spurious negative correlations caused by the sharing of common measurement errors by the initial scores and the gain scores. If the initial score is an over- or an under-estimate of the true score, the error included in the initial score will also carry over to the gain score. For example, if the observed score is X, the true score is Y, and the error is E, then $X = Y + E$. The change score is the difference between the two observed scores, $X_2 - X_1$. This can be rewritten $X_2 - (Y + E)$ which also contains the error E,

A second problem may be termed the unreliability-invalidity dilemma (Bereiter, 1963, p. 5). It stems from the fact that the higher the correlation between pre- and post-measurements, the lower the reliability of the difference scores. But the lower the correlation between them, the less they can be said to measure the same thing. This means that in situations where all cases show a similar amount of change, the differences between their change scores are not reliable because they may be entirely due to random disturbances. If the cases show widely varrying degrees of change, the second test may not be measuring the same thing as the first test when it was administered. For example, a test that measures arithmetic reasoning ability among young children may only measure computational accuracy when they grow older.

The two remaining problems cited in the literature are aspects of a single phenomenon. The first, called the 'ceiling effect', occurs if a maximum exists for a variable. When the value of the variable for a specific case falls near the ceiling during the first measurement, it cannot increase as much as for those cases whose values of the same variable were less (Duncan et al., 1961, p. 163). This becomes obvious when we recall that a student can increase his mark in an exam from forty percent to fifty percent much more easily than he can from eighty percent to ninety percent.

The alternative to the ceiling effect is called the 'zero boundary effect'. If the possible range of values is bounded by an absolute minimum, units with low scores will necessarily show less decline than units with initially high scores (Rummel, 1970, p. 237). Ceiling and zero boundary effects can be distinguished from regression effects although all three are concerned with movements of high and low scores. Regression effects deal with the average movement of high and low scores while the other two deal with the relative increase of high scores and relative decrease of low scores.

The most desirable measures of change would minimize or eliminate all the above-mentioned problems. Duncan et al, present four ways of measuring change - absolute, relative, positional, and deviational (1961, pp. 162-163). Absolute, or incremental change, as Rummel (1970) terms it, is the simple difference between scores at two points in time, $Y_2 - Y_1$. It is easy to compute but is subject to all the problems listed above: regression effects, ceiling and zero boundary effects, negative correlations between initial and change scores, and the unreliability-invalidity dilemma.

Relative change is the ratio of the change score, $Y_2 - Y_1$, to the initial score, Y_1. Percentage change is derived directly from this measure by multiplying by one hundred. It is also easy to compute but again is subject to all the problems involved in measuring change.

Positional and residual change are considerably more difficult to compute than the preceding two. Positional change is the difference in standard score position between the two measurements. It can be expresses as $Z_2 - Z_1$ where

$$Z_2 = \frac{Y_2 - \bar{Y}_2}{\sigma_2}$$

$$Z_1 = \frac{Y_1 - \bar{Y}_1}{\sigma_1}$$

This method partially compensates for the ceiling and zero boundary effects but is still plagued by the other difficulties in change measurement.

The fourth measure, deviational change, involves regressing values of later measures on those of the initial measurement. The resulting regression equation, $\hat{Y}_2 = A + BY_1$, provides expected values of the variable at the second time of measurement given its values at the time of the first measurement. Deviational change is the difference between the observed and expected acore, $Y_2 - \hat{Y}_2$. It is the change relative to the average change for all cases. It is considered by Rummel to be the only measure which surmounts all the problems of change measurement (1970, p. 239).

The operational definition of change in this study is deviational change. The 1961 values of each of the operational definitions of urban development, export base theory, sector theory, and growth pole theory are regressed on their corresponding 1951 values and the resulting deviation scores are used as measures of change.

The Model

Coleman recommends a multiple regression model to explain change in a dependent variable as a function of change in two or more independent variables. The model uses final and initial values of the dependent variables as well as initial and change values of the independent variables (Coleman, 1968, p. 443). The regression model takes the form

$$x_{1t} = a + b_1x_{10} + b_2x_{20} + b_{2\Delta}x_{2\Delta} + b_3x_{30} + b_{3\Delta}x_{3\Delta} + \ldots + b_ix_{io} + b_{i\Delta}x_{i\Delta} + \ldots + b_kx_{ko} + b_{k\Delta}x_{k\Delta}$$

where

b = the standardized partial regression coefficients;

x_{it} = the value of the dependent variable at the time of the second measurement;

x_{io} = the value of the independent variable i at the time of the earlier measurement;

x_{i_Δ} = the residual change of the independent variable i.

When change scores are defined in this way, as residuals of the regression of the later score on the initial score, they may be placed in the same equation as the initial scores because the two are not correlated (Bohrnstedt, 1968, p. 118).

The regression coefficients must be standardized to allow comparison because the unstandardized version is a function of the variation of the variable. Standardization eliminates the effect of variation and allows us to observe the relative contribution of each of the independent variables (Blalock, 1968, p. 189).

The model used in this thesis compares the relative explanatory powers of the three theories with respect to urban growth in Canada from 1951 to 1961. The 1951 values and residual values of each operational definition of the variables from the three theories are treated as independent variables in a multiple regression model. The dependent variable is 1961 average income while 1951 average income is considered another independent variable. The standardized partial regression coefficients indicate the relative contributions of each variable to total urban development.

Summary

The procedure in this chapter is to investigate and select the best operational definitions of crucial variables from each of the three theories. Operational definitions of change and of urban development are also selected and the unit of analysis is defined.

The units of analysis are the urban areas in Canada with populations in excess of 25,000 in 1961, with the exceptions of Corner Brook and Noranda-Rouyn, for which 1951 data are not available. Measurement of change is difficult but the method settled upon involves regression of 1961 values on corresponding 1951 values of each variable. The resulting residuals are the measures of change, more appropriately called 'residual change'.

The selection of an operational definition of export base theory is easy. A considerable proportion of the literature on export base theory is concerned with the problem of measuring the economic base of a region. The selected measure is the location quotient coefficient as calculated by Grossner (1970) and Crowley (1973). Two indicators of export base theory used here are the 1951 location quotient coefficient and the residual change in the location quotient coefficient, obtained from the residuals of the 1961 location quotient coefficient regressed on the 1951 location quotient coefficient.

The selection of indicators of sector theory is also straightforward. Definitions of manufacturing and service industries are well established. Walther Hoffman, who emphasized the importance of capital-goods industries in sector theory, defines capital-goods industries as the broad manufacturing sectors of metal and machinery, transportation equipment, and chemicals.

The six operational definitions of sector theory used in this thesis are:

1. percentage of the 1951 labour force engaged in manufacturing industries;
2. residual change in the percentage of the labour force engaged in manufacturing industries, obtained from the residuals when the percentage of the 1961 labour force engaged in manufacturing industries is regressed on the corresponding 1951 percentage;
3. percentage of the 1951 labour force engaged in service industries;
4. residual change in the percentage of the labour force engaged in service industries, obtained from the residuals when the percentage of the 1961 labour force engaged in service industries is regressed on the corresponding 1951 percentage;
5. percentage of the 1951 manufacturing labour force engaged in capital-goods industries;
6. residual change in the percentage of the manufacturing labour force engaged in capital-goods industries, obtained from the residuals when the percentage of the 1961 manufacturing labour force engaged in capital-goods industries is regressed on the correponding 1951 percentage;

The two variables selected as indicators of the strength of growth pole theory are agglomeration and entrepreneurial activity. The operational definitions of agglomeration are chosen to be the size of the labour force in 1951 and change in the size of the labour force. The operational definitions of entrepreneurial activities are the number of retail, wholesale, and service units in 1951 and the change in the number of retail, wholesale, and service units. The four operational definitions of variables significant to growth pole theory are:

1. 1951 labour force size;
2. residual change in labour force size, obtained from residuals when the 1961 labour force size is regressed on 1951 labour force size;
3. the number of retail, wholesale, and service units in 1951;
4. residual change in number of retail, wholesale, and service units, obtained from the residuals when the number of retail, wholesale, and service units in 1961 is regressed on the corresponding number in 1951.

The model used to compare the three theories is the multiple regression equation. The dependent variable is the operational definition of urban development, *viz*., the 1961 average wage of male wage earners. The independent variables are the indicators of 1951 variables and residual change in these variables from each of the three theories being tested. The usefulness of the three theories in explaining urban development in Canada from 1951 to 1961 depends on the relative sizes of the multiple regression coefficients of the independent variables.

Chapter 4

Data Analysis

This chapter describes the steps involved in the analysis of the data. Data analysis is defined as that part of the research process following the collection and processing of the data. After the data are transferred to data cards, the following steps occur:

1. Regression of 1961 variables on corresponding 1951 variables to obtain a matrix of residual values. These values are defined as measures of change.

2. The data are checked to make sure they meet the assumptions of regression analysis. This involves perusal of the distribution of residuals to ensure homoscedasticity and statistical independence. It also involves the construction of a correlation matrix for all 1951 and residual variables to detect the presence of collinearity among the data to be used as independent variables in the regression equation.

3. Factor analysis of the 1951 data and residual scores is done to obtain factor scores for those variables which exhibit collinearity. The factor scores are substituted for values of the corresponding collinear variables.

4. The dependent variable, 1961 average wage of male wage earners, is regressed on selected 1951 variables, residuals, and factor scores to obtain standardized partial regression coefficients for each independent variable.

5. Standardized partial regression coefficients are compared to determine the relative contributions of the independent variables to the variance of average wage of male wage earners in 1961. The theories associated with the variables can then be evaluated to determine their relative contributions to explaining urban economic development in Canada from 1951 to 1961.

Table 1 provides a summary of the 1951 and 1961 variables used as well as the concepts they represent and the theories from which the concepts are derived. Values of all variables are extracted from 1951 and 1961 census publications, except for location quotients which are calculated by Grossner (1970).

The 1961 values are regressed on corresponding 1951 values to obtain expected values for 1961. The residuals, which are differences between observed and expected 1961 scores, represent deviational change over the decade, 1951-1961.

Table 1

Variables for Which Data are Collected

Theory	Concept	Variables
Urban development	Regional income	Average earnings of male wage earners 1951, 1961
Sector theory	Percentage of labour force in manufacturing industries	Percentage of labour force in manufacturing industries, 1951, 1961
	Percentage of labour force in service industries	Percentage of labour force in service industries, 1951, 1961
	Percentage of manufacturing labour force in capital-goods industries	Percentage of manufacturing labour force in capital-goods industries, 1951, 1961
Export base theory	Relative proportion of labour force in export industries	Location quotient coefficient, 1951, 1961
Growth Pole Theory	Agglomeration	Labour force size, 1951, 1961
	Entrepreneurial activity	Number of retail, wholesale, and service units, 1951, 1961

The question arises whether the data meet the requirements of regression analysis. According to Wonnacott and Wonnacott (1970, p. 15) the assumptions in regression analysis are that in the equation $Y = A + BX + e$:

1. the probability functions, $p(y_i / x_i)$, have the same variance, σ, for all x_i;
2. the probability functions, $p(y_i / x_i)$, have the means, $E(y_i)$, lying on a straight line known as the true regression line;
3. values of Y_i are statistically independent;
4. values of x_i are independent of A, B, and σ;
5. the distribution of e is normal and its values are independent of X.

To ensure that the data can be entered into a regression equation, values of Y are plotted against values of X. The plots reveal that they could be represented by a straight line. Additional examination of the residuals is necessary because they are used as input into the final multiple regression equation. The spread of deviations must be constant about the regression line and about the values of the independent variable as well as about the observed and expected values of the dependent variable (Draper and Smith, 1966, p. 122). The plots reveal a satisfactory distribution in all cases.

Table 2

Simple Regression of 1961 Variables on Corresponding 1951 Variables

	Variable*					
	1	2	3	4	5	6
Proportion of variance reduced	0.995	0.797	0.970	0.972	0.801	0.994
Correlation coefficient	0.997	0.893	0.985	0.986	0.895	0.997
Standard error of estimate	12057	3.957	2.461	2.280	9.319	450
F	8455	173	1423	1507	177	6893
a**	-1544.8	0.879	0.628	16.35	3.341	-28.13
b**	1.421	1.235	0.793	0.816	0.814	1.235
Standardized b	0.997	0.893	0.985	0.986	0.895	0.997

*1. Labour force size

2. Location quotient.

3. Percentage of labour force in manufacturing.

4. Percentage of labour force in service industries.

5. Percentage of manufacturing labour force in capital-goods industries.

6. Number of retail, wholesale, and service establishments.

**$Y = a + bX$

The conclusion derived from observing the plots is supported by data in Table 2. The F-tests are all significant at the 0.001 level. They confirm that a substantial proportion of the sums of squared deviations about the mean is explained by the regression line.

The regression line reduces the greatest proportion of variance of the labour force size and number of establishments while reducing proportion of the variance of the location quotient least. The correlation coefficients can be interpreted as stability coefficients of our variables between 1951 and 1961. Labour force size, number of establishments, and percentages of the labour force in manufacturing and in service industries appear to be the most stable variables while the location quotient and proportion of manufacturing labour force in capital-goods industries are less stable over the ten year period. The standardized regression coefficients are equivalent to correlation coefficients in the simple two variable case. Table 2 indicates that all the 1951 variables act as excellent predictors of corresponding 1961 variables.

The desired set of data is now available for use in the multiple regression equation, but multicollinearity is still possible. The occurrence of multicollinearity causes the model to lose one dimension so that the estimation procedure for obtaining the partial regression coefficient becomes sensitive to random errors (Wonnacott and Wonnacott, 1970, pp. 60-61). A product-moment correlation analysis is performed on the set of independent variables to determine if any pairs are collinear. Collinearity is arbitrarily defined as occurring when the product-moment correlation coefficient of two variables is greater than 0.7. The independent variables are:

1. average wage of 1951 male wage earners;
2. 1951 labour force size;
3. 1951 location quotient;
4. percentage of 1951 labour force in manufacturing industries;
5. percentage of 1951 labour force in service industries;
6. percentage of 1951 manufacturing labour force in capital-goods industries;
7. number of retail, wholesale, and service establishments in 1951;
8. residual change in labour force size, 1951-1961;
9. residual change in location quotient, 1951-1961;
10. residual change in percentage of the labour force engaged in manufacturing activities, 1951-1961;
11. residual change in percentage of the labour force engaged in service industries, 1951-1961.
12. residual change in percentage of the manufacturing labour force engaged in capital-goods industries, 1951-1961;

13. residual change in the number of retail, wholesale, and service units, 1951-1961.

The following pairs had correlation coefficients greater than 0.7:

- 1951 labour force size and number of retail, wholesale, and service units in 1951 (0.99);
- change in labour force size and change in number of retail, wholesale, and service units (0.89);
- 1951 percentages of labour force engaged in manufacturing and in service industries (-0.86).

It is interesting to note that collinear variables represent the same theories. The first two pairs derive from growth pole theory while the third pair are both indicators of sector theory. The variables were originally selected to represent separate components of each theory but it is understandable that separate components may be correlated. Lack of high correlation, however, cannot be construed conversely to mean that the variables do not represent the same theory.

The variables were factor analyzed and values of collinear variables were replaced by their factor scores. A common factor model was used with initial communality estimates equal to squared multiple correlations. Four factors were obtained with eigenvalues greater than 0.1 and these were rotated obliquely to obtain optimum identification of variable clusters independent of the locations of other clusters.

The first factor loaded highly on percentage of the 1951 labout force in manufacturing industries and percentage of the 1951 labour force in service industries. The second factor loaded heavily on residual change of labour force size and residual change of retail, wholesale, and service units. The third factor had high loadings on 1951 labour force size and on number of retail, wholesale, and service units in 1951.

Collinear variables were replaced by the derived factor scores and regression analysis was applied. The regression equation consisted of:

1. the dependent variable, 1961 average wage of male wage earners, and ten independent variables:
2. average wage of 1951 male wage earners;
3. 1951 location quotient;
4. percentage of 1951 manufacturing labour force in capital-goods industries;
5. residual change in location quotient, 1951-1961;
6. residual change in percentage of the labour force engaged in manufacturing industries;
7. residual change in percentage of the labour force engaged in service industries;
8. residual change in percentage of the manufacturing labour force engaged in capital-goods industries;

9. factor score of percentages of 1951 labour force in manufacturing, and in service industries;
10. factor score of residual changes in labour force and in number of retail, wholesale, and service units, 1951-1961;
11. factor score of 1951 labour force size and number of retail, wholesale, and service units in 1951.

Table 3 shows the simple correlations among the variables used in the regression equation. The correlation between 1951 and 1961 average wages of male wage earners is 0.85, the highest correlation coefficient in the table. The coefficient indicates collinearity between the two variables but this does not present a problem when it occurs between a dependent and an independent variable. It indicates a high rate of stability in wages which reflects that those factors influencing wages of male wage earners in 1961 were also at work in 1951 and probably in earlier decades as well. No other dependent variable correlates highly with average wage of 1961 male wage earners; the second highest correlate is proportion of the 1951 manufacturing labour force in capital-goods industries, with a coefficient of 0.35, but as we shall see, most of this correlation is partialled out when we control for other variables, especially 1951 average wage, with which it has a correlation coefficient of 0.50.

Table 3

Simple Correlation Matrix of Eleven Variables Used in Regression Equations

Variable*

	1	2	3	4	5	6	7	8	9	10
2	0.85									
3	0.01	0.03								
4	0.35	0.50	0.27							
5	0.30	0.34	-0.00	0.43						
6	-0.13	-0.14	-0.07	-0.06	-0.13					
7	-0.00	0.01	0.02	0.15	0.04	-0.68				
8	0.12	0.13	0.16	0.00	-0.06	0.24	-0.19			
9	-0.15	-0.39	-0.20	-0.55	-0.43	0.12	-0.11	-0.01		
10	0.28	0.31	0.00	0.08	0.02	0.03	-0.12	0.16	0.02	
11	0.32	0.17	-0.13	-0.07	0.01	0.02	-0.02	-0.01	0.16	0.08

*1. 1961 average wage
2. 1951 average wage
3. 1951 location quotient
4. 1951 % capital goods
5. Change in location quotient
6. Change in % manufacturing
7. Change in % service
8. Change in % capital goods
9. 1951 % manufacturing and service
10. Change in labour force and units
11. 1951 labour force and units

The highest correlation coefficient among the independent variables is -0.68 which occurs between change in percentage of the labour force engaged in service industries and change in percentage engaged in manufacturing industries. We would expect a relatively high negative correlation between these two since a proportionate gain in one would require a proportionate decline in the other or in primary industries.

The second highest correlation coefficient among the independent variables, -0.55, occurs between the percentage of 1951 manufacturing labour force engaged in capital-goods industries and the factor score summarizing the percentage of the 1951 labour force engaged in manufacturing and in service industries. A statistical explanation does not account for this relationship as it did in the previous case. Apparently, when a large fraction of the labour force is engaged in either manufacturing or services, a small fraction of the manufacturing labour force is engaged in capital-goods industries. We may pursue this relationship further by checking the correlation coefficients of proportion of the manufacturing labour force engaged in capital-goods industries with each of the percentages of the total labour force engaged in manufacturing (0.28) and in service industries (-0.35). Neither coefficient suggests that a factor score of the latter two would correlate -0.55 with percentage of manufacturing labour force producing capital goods although the percentage in service industries is most similarly correlated.

The remainder of the coefficients of the independent variables are smaller than those mentioned in the preceding three paragraphs. They range in value from 0.43 (proportion of 1951 manufacturing labour force engaged in capital-goods industries and residual change in location quotient) to -0.43 (residual change in location quotient and factor score summarizing percentages of the 1951 labour force in manufacturing and service industries).

Examination of the contribution of each of the independent variables to variation in the dependent variable is facilitated by a stepwise multiple regression model. According to this procedure the X (independent variable) most highly correlated with the Y (dependent variable) is first entered into the regression equation. The second step involves entering into the regression equation one of the remaining X variables which has the highest partial correlation coefficient with Y controlling for the one already in the equation. In the third step, a sequential F-test is used to determine the contribution the first X variable would make if the second one were entered first. Sometimes its contribution is considerably reduced by the subsequent introduction of another variable into the regression equation. The remaining steps involve a repetition of the second and third steps but with a difference in the selection process. All variables already in the equation are partialled out to determine which of the remaining variables have the highest partial correlation with the dependent variable (Draper and Smith, 1966, pp. 171-172).

Table 4 presents the regression coefficients of the six independent variables which contribute to the reduction of unexplained variance when it has been entered into the equation. It also displays the proportionate reduction and cumulative proportionate reduction of unexplained variance, for each step in the stepwise regression analysis. The best predictor of 1961 average wage of male wage earners is the 1951 average wage of male wage earners. This does not detract from the contributions of the three theories to explanation of the variation in 1961 average wage because the 1951 wage levels are presumably a function of prior forces exerted by the three theories. However, in this study, 1951 average wage has been included only to provide a stability coefficient. It is supposed to reflect the influence of economic forces of the three theories previous to 1951 rather than being an indicator of a single theory.

Table 4 also reveals that very little reduction in unexplained variance occurs after the fourth step. This suggests that the only variables worth retaining in the regression equation are average wage of 1951 male wage earners (X_2), the factor score summarizing percentages of the 1951 labour force in manufacturing and service industries (X_9), another factor score summarizing the size of the 1951 labour force and the number of retail, wholesale, and service units in 1951 (X_{11}), and the residual change in the location quotient (X_5). These four explain 78.9 percent of the variation of the dependent variable while inclusion of the remaining two independent

Table 4

Regression Coefficients of Six Variables and Contributions at Each Step to Reduction of SS of 1961 Wage Variance in Stepwise Regression

		Variables*							
Step	Intercept	2	9	11	5	3	6	Prop. Variance Reduced	Cumulative Prop. Variance Reduced
1	298	1.55						0.728	0.728
2	-34	1.69	83.12					0.037	0.765
3	125	1.62	68.05	54.24				0.017	0.783
4	198	1.59	81.22	52.82	9.21			0.007	0.789
5	180	1.59	85.39	54.41	9.63	2.77		0.002	0.791
6	188	1.59	85.62	54.63	9.47	2.71	-2.97	0.001	0.792

2. 1951 average wage
9. 1951 % manufacturing and service
11. 1951 labour force and units

5. Change in location quotient
3. 1951 location quotient
6. Change in % manufacturing

Table 5

Standardized Regression Coefficients of Six Variables at each Step of Stepwise Regression Analysis

	Variables*					
Step	2	9	11	5	3	6
1	0.85					
2	0.93	0.21				
3	0.90	0.17	0.13			
4	0.88	0.20	0.13	0.09		
5	0.88	0.21	0.14	0.10	0.04	
6	0.88	0.21	0.14	0.09	0.04	-0.02

2. 1951 average wage
9. 1951 % manufacturing and service
11. 1951 labour force and units
5. Change in location quotient
3. 1951 location quotient
6. Change in % manufacturing

variables only explains an additional 0.3 percent of the variation. Restricting the equation to these four variables provides the following regression equation:

$$Y = 198 + 1.59X_2 + 81.22X_9 + 52.82X_{11} + 9.21X_5$$

Table 5 presents the standardized regression coefficients corresponding to the unstandardized coefficients displayed in Table 4. Standardized regression coefficients allow the relative contributions of the independent variables to variation in the dependent variable to be compared. The coefficients can be interpreted as the fraction of the standard deviation of the dependent variable for which the independent variable is directly responsible assuming that all other independent variables remain constant. The inefficacy of variables five and six becomes more obvious in Table 5 where their standardized regression coefficients are seen to approach zero. The first four independent variables contribute to the standardized regression equation as follows:

$$Y = 0.88X_2 + 0.20X_9 + 0.13X_{11} + 0.09X_5$$

Table 5 presents a possible solution to the problem posed in this thesis: what are the relative contributions of each theory to urban development in Canada from 1951 to 1961? Variables nine, eleven, and five each represent separate theories. Variable eleven, which is responsible for about fifty percent more variation in the dependent variable than variable five, represents growth pole theory. Variable nine, which contributes about fifty percent more than variable eleven and about one hundred percent more than variable five, represents sector theory. Apparently sector theory is most powerful of the three with growth pole theory second and export base theory third.

A qualification must be made before this summarizing statement can be accepted: the assumptions of multiple regression analysis must be valid in the present study. The assumptions as outlined by Darlington (1970, p. 370) are:

1. all variables which might affect the dependent variable are either included in the regression equation or are not correlated with the variables which are included;
2. all relationships between the variables are linear and non-interactive;
3. values of the dependent variable have no effect on values of the independent variables.

Only when these assumptions are met can the standardized partial regression coefficients be interpreted as measures of direct causal effects of the independent variables on the dependent variable.

The third assumption is sound because values of the dependent variable, 1961 average wage of male wage earners, occur subsequently to those of all the independent variables which pertain to 1951 and to change from 1951 to 1961. The second assumption is a reasonable one given that no knowledge of curvilinear relationships exists. The assumption of linearity is the parsimonious one to adopt until disproven or until it seems logical to discard it.

The first assumption is the one that may cause misgiving. No authority on regional or urban development states that these three theories explain all of urban development. A number of social and demographic factors, as well as economic and political influences beyond the scope of each of the theories, also play essential roles in regional and urban development. Apprehension is increased when it is recognized that the concepts selected from the theories in this study are crucial but not comprehensive. Growth pole theory recognizes other factors besides agglomaration and entrepreneurial activity as factors of growth. Sector theory concedes development to be more complex than simple proportions of the labour force in designated industries indicates. Export base theory recognizes that external demand is an essential complement to a large export base although it is difficult to measure. To assume that external variables are not correlated with any included in the regression equation may be somewhat heroic.

If we wish to confine our analysis to prediction rather than causal analysis, a measure of 'usefulness' may be of greater interest than standardized regression coefficients. Usefulness has been defined as the amount by which the squared multiple correlation would diminish if the independent variable were removed from the regression equation and the remaining variables appropriately reweighted (Darlington, 1970, p. 362). The usefulness coefficient can be computed from the following formula:

$$U_j = \frac{P_j^2}{1 - P_j^2} (1 - R^2)$$

where:

U_j = the usefulness of variable j;

P_j = the partial correlation between Y and X_j partialling out the other predictor variables;

R^2 = the squared multiple correlation (Darlington, 1970, p. 371).

Since the squared multiple correlation is constant for each predictor variable in a set, the measure of usefulness is a transform of the partial correlation between the dependent and independent variable, controlling for all other independent variables.

Table 6 reveals that similar ranking of the variables obtains using standardized partial regression coefficients, partial correlations, and the measure of usefulness. The table also provides an alternative criterion of ranking variables by importance, viz., the validity coefficient or coefficient of determination. It is equivalent to r^2 and can be interpreted as the amount of unexplained variation in the dependent variable reduced by knowledge of the independent variable. It has the advantage over U and β of being constant regardless of the number and choice of other independent variables included in the regression equation. It has the disadvantage of being a measure of association rather than of causality or prediction.

Table 6

Measures of Importance of Ten Variables in the Regression Equation

	β	Partial Correlation	U^*	r	r^2
1951 average wage	0.88	0.82	0.43	0.85	0.72
1951 % manufacturing, service	0.21	0.33	0.03	-0.15	0.02
1951 labour force and units	0.14	0.28	0.02	0.32	0.10
Change in location quotient	0.10	0.18	0.01	0.30	0.09
1951 location quotient	0.04	0.09	0.00	0.01	0.00
Change in % manufacturing	-0.02	-0.03	0.00	-0.13	0.02
Change in % capital goods	0.02	0.04	0.00	0.12	0.01
1951 % capital goods	-0.02	-0.03	0.00	0.35	0.12
Change in labour force and units	-0.01	-0.02	0.00	0.28	0.08
Change in % service	0.00	0.00	0.00	-0.00	0.00

$$^*U = \frac{P^2}{1-P^2}(1-R^2)$$

where P = the partial correlation coefficient

and R = the multiple correlation coefficient

In Table 6 we note that four variables with less predictive power than the factor score summarizing percentages of the 1951 labour force in manufacturing and service industries have higher validity coefficients. The four are the factor score summarizing percentages of the 1951 labour force size and the number of retail, wholesale, and service units in 1951, residual change in the location quotient, percentage of the 1951 manufacturing labour force engaged in capital-goods industries, and factor scores summarizing residual changes in labour force size and number of retail, wholesale, and service units.

Perusal of Table 3 reveals that these four variables have simple correlation coefficients of 0.17, 0.34, 0.50, and 0.31 with average wage of 1951 male wage earners. Possibly the partial correlation coefficients of these four variables are low because of their association with 1951 average wage. It seems theoretically incorrect to partial out their association with 1951 average wage to determine their relationship with 1961 average wage. The former is the same variable as the latter, but measured at a different point in time. Variables which are associated with income can be expected to correlate with it at both points in time, given a stable political, social, and economic environment. Thus the inclusion of 1951 average wage in the regression equation may suppress relationships between the dependent and some of the other independent variables as measured by standardized partial regression coefficients and by the usefulness criterion.

Table 7

Regression Coefficients of Nine Variables and Contribution of each Step to Reduction of SS of 1961 Income Variance in Stepwise Regression

		Variable*										
Step	Intercept	4	11	10	5	8	6	7	3	9	Prop. Variance Reduced	Cumltve. Prop. Variance Reduced
1	3600	6.05									0.123	0.123
2	3586	6.45	136								0.120	0.243
3	3598	6.13	128	75.7							0.051	0.294
4	3644	4.82	125	76.8	17.4						0.025	0.319
5	3646	4.79	126	71.4	18.2	4.45					0.011	0.329
6	3646	4.77	127	71.1	16.7	5.76	-21.5				0.017	0.346
7	3630	5.21	128	65.3	14.9	5.62	-38.3	-26.9			0.012	0.358
8	3641	5.67	125	64.0	13.7	6.20	-40.7	-28.7	-4.82		0.005	0.363
9	3628	6.00	123	63.2	14.7	6.26	-41.3	-28.9	-4.66	17.8	0.001	0.364

*4. 1951 % capital goods

11. 1951 labour force and units

10. Change in labour force and units

5. Change in location quotient

8. Change in % capital goods

6. Change in % manufacturing

7. Change in % services

3. 1951 location quotient

9. 1951 % manufacturing and service

Table 8

Standardized Regression Coefficients of Nine Variables at each Step of Stepwise Regression Analysis

	Variable*								
Step	4	11	10	5	8	6	7	3	9
1	0.35								
2	0.37	0.35							
3	0.36	0.33	0.23						
4	0.28	0.32	0.23	0.17					
5	0.28	0.32	0.22	0.18	0.10				
6	0.28	0.32	0.21	0.17	0.14	-0.13			
7	0.30	0.33	0.20	0.15	0.13	-0.24	-0.15		
8	0.33	0.32	0.19	0.14	0.15	-0.25	-0.16	-0.08	
9	0.35	0.31	0.19	0.15	0.15	-0.26	-0.17	-0.07	0.04

*4. 1951 % capital goods
11. 1951 labour force and units
10. Change in labour force and units
5. Change in location quotient
8. Change in % capital goods
6. Change in % manufacturing
7. Change in % services
3. 1951 location quotient
9. 1951 % manufacturing and service

Tables 7 and 8 provide step-by-step unstandardized and standardized regression coefficients with average wage of 1951 male wage earners removed from the regression equation. Removal of the variable results in a radical rearrangement of variables. The variable which is deemed most important behind 1951 average wage is rated least important when the 1951 average wage is removed from the equation. Another variable which loses considerable importance is the 1951 location quotient. The two variables which gain in importance are: percentage of the 1951 manufacturing labour force engaged in capital-goods industries, which climbs from eighth to first, and the factor score summarizing labour force size and number of retail, wholesale, and service units in 1951, which climbs from ninth to third in importance.

The column dealing with proportion of reduced variance in Table 7 reveals that most of the variables play a more significant part in explaining variation in 1961 average wage with the partialling effect of 1951 average wage removed. They explain 36.4 percent of the variation in 1961 average wage whereas under the suppressive effect of 1951 average wage in Table 4 they explained only 6.4 percent (79.2 - 72.8). Every variable increased its contribution to unexplained variance of 1961 average wage with the exception of the factor score summarizing the percentages of the 1951 labour force engaged in manufacturing and service industries.

Table 9

Measures of Importance of Nine Variables in the Regression Equation

	β	Partial Correlation	U^*	r	r^2
1951 % capital goods	0.35	0.31	0.07	0.35	0.12
1951 labour force and units	0.31	0.36	0.10	0.32	0.10
Change in labour force and units	0.19	0.23	0.04	0.28	0.08
Change in location quotient	0.15	0.15	0.02	0.30	0.09
Change in % capital goods	0.15	0.17	0.02	0.12	0.01
Change in % manufacturing	-0.26	-0.22	0.03	-0.13	0.02
Change in % service	-0.17	-0.15	0.02	-0.00	0.00
1951 location quotient	-0.07	-0.09	0.01	0.01	0.00
1951 % manufacturing and service	0.04	0.04	0.00	-0.15	0.02

$$^*U = \frac{P^2}{1-P^2}\,(1-R^2)$$

where P = the partial correlation coefficient

and R = the multiple correlation coefficient

Table 9 presents the criteria by which the relative importance of the nine independent variables can be discerned. If we accept the assumptions inherent in the regression model, especially the one stating that all other causal variables are uncorrelated with those in the regression equation, then the standardized regression coefficients are the best criteria. Their values suggest that percentage of the 1951 manufacturing labour force engaged in capital-goods industries is the most important cause of variation in 1961 average wage. The factor score summarizing 1951 labour force size and number of retail, wholesale, and service units ranks second while residual change in percentage of the labour force engaged in manufacturing ranks third. A disturbing feature of this column is the negative sign of the last-mentioned variable. This can be interpreted to mean that an increase in percentage of the labour force in manufacturing industries results in a decline in the average wage, which contradicts sector theory rather than representing it. The negative value of the variable indicates that the theory does not explain urban development in spite of the fact that the most powerful explanatory variable in the regression equation, percentage of 1951 manufacturing labour force in capital-goods industries, is also drawn from sector theory. Apparently one aspect of sector theory does a relatively good job of explaining urban development while another explains the inverse - urban attrition.

In spite of the rearrangement of individual variables, Table 9 supports the ranking of the theories found in Table 6 in terms of causal analysis, β, and predictive usefulness, U. Sector theory is again ranked first, growth pole theory second, and export base theory third.

Dropping 1951 average wage from the set of predictor variables reduces the amount of cumulatively explained variance from 0.792 to 0.364. This reduction suggests that the latter model is inadequate for explaining variation in 1961 average wage as 1951 average wage obviously plays a key role in determining wages ten years later. An alternative model that should be considered regresses residual change in average wage on the set of independent variables. Residual change in average wage does not depend on 1951 average wage as does 1961 average wage, being statistically independent. Thus it is possible that 1951 variables and change variables drawn from the three theories being investigated may explain a larger proportion of the variation in residual change of average wage than in 1961 average wage. The change scores of the independent variables may or may not be considered causes of change in the dependent variable which is also a change score for the same period of time. However the period of change is long enough, ten years, that the relationship between them may be causal.

Table 10

Regression Coefficients of Six Variables and Contribution of Each Step to Reduction of SS of Change in Wage in Stepwise Regression

		Variable*							Cumltve. Prop.
Step	Intercept	9	11	5	3	6	8	Prop. Variance Reduced	Variance Reduced
1	-0.00	70.6						0.115	0.115
2	-0.00	60.9	58.2					0.079	0.194
3	-0.01	77.8	55.0	9.66				0.028	0.222
4	-15.6	81.9	56.6	10.1	2.75			0.007	0.229
5	-14.8	82.5	56.7	9.88	2.69	-3.26		0.001	0.230
6	-13.3	82.7	56.5	10.0	2.41	-4.28	1.10	0.002	0.232

*9. 1951 % manufacturing and service
11. 1951 labour force and units
5. Change in location quotient
3. 1951 location quotient
6. Change in % manufacturing
8. Change in % capital goods

Table 11

Standardized Regression Coefficients of Six Variables at Each Step of Stepwise Regression Analysis

	Variable*					
	9	11	5	3	6	8
1	0.34					
2	0.29	0.29				
3	0.37	0.27	0.18			
4	0.39	0.28	0.19	0.08		
5	0.40	0.28	0.19	0.08	-0.04	
6	0.40	0.28	0.19	0.07	-0.05	0.05

*9. 1951 % manufacturing and service

11. 1951 labour force and units

5. Change in location quotient

3. 1951 location quotient

6. Change in % manufacturing

8. Change in % capital goods

Tables 10 and 11 present the unstandardized and standardized partial regression coefficients, respectively, of the regression equation with residual change in average wage as the dependent variable. Table 10 also reflects the limited powers of the independent variables in reducing unexplained variance in the residual change in average wage. Altogether they account for only 23.2 percent of the total variation of the dependent variable. The first three variables, the factor score summarizing percentages of the 1951 labour force engaged in manufacturing and service industries (X_9), the factor score summarizing labour force size and number of retail, wholesale, and service units in 1951 (X_{11}), and the residual change in the location quotient (X_5), account for 22.2 percent of the total variation of the dependent variable. The most parsimonious model would include these three variables, and, written in standardized notation, would be:

$$Y = 0.37X_9 + 0.27X_{11} + 0.18X_5$$

Once again the best predictor comes from sector theory, the second best from growth pole theory, and the third best from export base theory. Table 12 demonstrates that with one minor exception the variables are ranked in the same way using beta weights or Darlington's criterion of usefulness. The validity coefficient, which is a measure of association rather than causality or prediction, ranks the first two variables equivalent to that of β and U, but places percentage of 1951 manufacturing labour force in capital-goods industries third.

Table 12

Measures of Importance of Nine Variables in the Regression Equation

	β	Partial Correlation	U*	r	r^2
1951 % manufacturing and service	0.39	0.33	0.09	0.34	0.12
1951 labour force and units	0.28	0.30	0.08	0.33	0.11
Change in location quotient	0.19	0.18	0.03	0.03	0.00
1951 location quotient	0.08	0.08	0.00	-0.03	0.00
Change in % manufacturing	-0.05	-0.04	0.00	-0.01	0.00
Change in % capital goods	0.05	0.05	0.00	0.03	0.00
1951 % capital goods	-0.02	-0.01	0.00	-0.14	0.02
Change in labour force and units	-0.01	-0.01	0.00	0.03	0.00
Change in % services	-0.01	-0.01	0.00	-0.02	0.00

$$^{*}U = \frac{P^2 \cdot (1-R^2)}{1-P^2}$$

where P = the partial correlation coefficient

and R = the multiple correlation coefficient

Tables 10, 11, and 12 refer to the regression model which takes residual change in average wage as the dependent variable. It was chosen as an alternative to the model using 1961 average wage as the dependent variable and excluding 1951 average wage as an independent variable. However the latter model, which takes 1961 average wage of male wage earners as the dependent variable, provides a greater reduction in unexplained variance of the dependent variable and is considered the best of the three models that have been entertained. It eliminates 1951 average wage, which was suppressing the regression of 1961 average wage on other independent variables and it explains more variation of the dependent variable than the model taking residual change in average wage as the dependent variable.

A disadvantage of these three models for comparing three theories is the fact that they compare variables rather than theories. A method that allows more direct comparison of the theories would involve separate regression of the dependent variable on the sets of variables from each theory. The cumulative proportionate reduction in unexplained variance by each set of variables would indicate the relative explanatory powers of the three theories.

Table 13

Measures of Importance of Variables From Each Theory

	β	Proportion Variance Reduced	Cumulative Proportion Variance Reduced
Sector Theory			
1951 % capital goods	0.41	0.123	0.123
Change in % capital goods	0.15	0.016	0.139
Change in % manufacturing	-0.30	0.019	0.158
Change in % service	-0.23	0.026	0.184
1951 % manufacturing and service	0.09	0.005	0.189
Growth Pole Theory			
1951 labour force and units	0.30	0.104	0.104
Change in labour force and units	0.26	0.065	0.169
Export Base Theory			
Change in location quotient	0.30	0.092	0.092
1951 location quotient	0.01	0.000	0.092

Table 13 indicates that three separate regression equations each containing variables from a different theory rank the theories in the same order as did a single regression model containing all the variables. Variables from sector theory account for 18.9 percent of total variation in 1961 average wage, variables from growth pole theory account for 16.9 percent, while those from export base theory account for only 9.2 percent.

Sector theory has an advantage over the other two in that five variables are used in the regression equation. while only two are used from each of the other theories. Additional variables would almost certainly increase the cumulative proportionate reduction in variance of the dependent variable. A fairer comparison would allow an equal number of variables in each regression equation. The two variables from growth pole theory provide more cumulative proportionate reduction in variance of the dependent variable than do the first two from sector theory, but it is possible that a different combination of two variables from sector theory might provide more cumulative proportionate reduction in variance.

Four regression equations were derived to determine the maximum cumulative proportionate reduction in variance attained combining percentage of the 1951 manufacturing labour force engaged in capital-goods industries with each of the other variables from sector theory.

Table 14

Measures of Importance of Variables from Sector Theory and Growth Pole Theory

Sector Theory	β	Proportionate Reduced Variance	Cumulative Proportionate Reduced Variance
1951 % capital goods	0.34	0.123	0.123
Change in % manufacturing	-0.11	0.011	0.134
1951 % capital goods	0.34	0.123	0.123
Change in % services	-0.05	0.003	0.126
1951 % capital goods	0.34	0.123	0.123
Change in % capital goods	0.12	0.016	0.138
1951 % capital goods	0.34	0.123	0.123
1951 % manufacturing and service	0.06	0.002	0.125
Growth Pole Theory			
1951 labour force and units	0.30	0.104	0.104
Change in labour force and units	0.26	0.065	0.169

The results are displayed in Table 14. The most powerful combination of variables in sector theory, percentage of the 1951 manufacturing labour force engaged in capital-goods industries, and residual change in this same percentage, reduces the unexplained variance of 1961 average wage by 13.8 percent which is less than the 16.9 percent reduced by the two growth pole theory variables.

Table 14 indicates that when two variables are taken from each theory, growth pole theory best explains urban development in Canada from 1951 to 1961. It is followed closely by sector theory while export base theory lags behind. The two variables from growth pole theory account for 16.9 percent of the cumulative proportionate reduction in variance while the best combination of two sector theory variables can account for only 13.8 percent of this variance. The best single predictor is percentage of the 1951 manufacturing labour force engaged in capital-goods industries which derives from sector theory. Export base theory cannot be dismissed because the most powerful change variable is residual change in the location quotient. It measures change from 1951 to 1961 and derives from export base theory. An explanation of economic conditions in Canada in 1961 which looks to change in the preceding decade would necessarily give some credance to export base theory.

A more adequate appraisal of the contributions of the three theories to urban development in Canada would entail recruitment of more variables from growth pole theory and export base theory. After constructing factor scores of collinear variables, we were left with fewer variables to represent these two theories. The preceding tables indicate that sector theory produces more cumulative proportionate reduction in variance of the dependent variable on the strength of number of variables rather than because individual variables can significantly reduce the amount of unexplained variance.

A search for additional variables for export base theory seems futile because the location quotient coefficient _per se_ is crucial to that theory. Other variables cannot save the theory if the location quotient coefficient fails. If the 1951 location quotient coefficient and residual change in the location quotient coefficient do not explain urban development, as measured by change in average wage of male wage earners, no other variable can redeem export base theory because these variables measure a crucial aspect of the theory - they measure the size of the export base.

It was consequently decided to discard export base theory and concentrate on selecting three more variables to represent growth pole theory. The three selected variables include an alternative measure of agglomeration (migration impact), and two measures of the profit-making ability of firms (1951 average value added and residual change in average value added).

Migration Impact

It can be argued that change in labour force size is the result of the interplay of two sets of forces: economic and demographic. The economic view is that labour force size changes as a function of the changing number of job opportunities. The demographic view emphasizes that changing numbers of persons available for work alters the size of the work force. From the economic perspective, the best indicator of agglomeration is the change in the labour force size. From the demographic perspective, change in labour force size depends on change in the number of persons of labour force age, which is a function of:

1. net migration;
2. changes in sex- and age-specific labour force participation rates;
3. natural increase.

Female participation rates increased over the decade in Canada but male participation rates remained fairly constant (Denton and Ostry, 1967, pp. 26-27; Ostry, 1968, p. 13). Data on participation rates are not available for individual urban areas but under the assumption that these rates are equal for all cities, variations in the labour force growth can be attributed to variations in net migration and natural increase. Migration is the component which is sensitive to agglomeration while natural increase is a reflection of the demographic history of the urban area.

If it is argued that expansion of the labour force can be caused by availability of persons of labour force age, then net migration of persons of labour force age is a better indicator of agglomeration than change in labour force size because change in labour force size also reflects natural increase. The measure of migration must necessarily be restricted to that of members of the labour force to be a valid indicator of agglomeration because total migration figures also reflect age structure and family size of migrants as well as including migrants of retirement age.

Measures of net migration of persons in the labour force are calculated for each urban area to determine whether these are useful in explaining variation in 1961 average wage of male wage earners. The method used is called the forward lifetable survival rate method, adapted here to measure migrants of working age rather than total migration. This method involves applying the following formula to each five- and ten-year age group in each city. The span of the age groups was selected to coincide with published classifications of 1961 age groups in the labour force. The formula which obtains net number of migrants of working age in each city is:

$$M = \sum_{i=1}^{9} \left(LF_{i+10}^{1961} - P_{i}^{1951} \cdot \frac{{}_{n}L_{i+10}}{{}_{n}L_{i}} \cdot R_{i+10}^{1961} \right) \frac{{}_{n}L_{i+5}}{{}_{n}L_{i+10}}$$

where:

M = net migration of persons in the labour force;

i = age groups 5-9, 10-14, 15-24, 25-34, 35-44, 45-49, 50-54, 55-59, 60-64;

n = 5 or 10 (span of the age group);

LF_{i+10}^{1961} = 1961 labour force population of age group i+10;

P_{i}^{1951} = 1951 population of age group i;

$\frac{{}_{n}L_{i+10}}{{}_{n}L_{i}}$ = ten year life table survival rate for age group i;

R_{i+10}^{1961} = labour force participation rate for age group i+10 in 1961;

$\frac{{}_{n}L_{i+5}}{{}_{n}L_{i+10}}$ = inverse life table survival rate for migrants age i+10;

The last term attempts to capture those in- and out-migrants who die before 1961. Without it, the forward life table survival rate method provides an underestimation of net migration in cases of both net in- and out-migration. In cases of net in-migration the population at risk of mortality increases above the initial level. The procedure does not account for those persons who migrated into the region and died. Thus it underestimates both net migration and mortality. In cases of out-migration the population at risk of mortality falls below the initial level. The procedure credits the loss of those individuals who emigrate before dying to mortality when it should credit the loss to out-migration. Thus it overestimates mortality and underestimates out-migration.

One suggested method of circumventing this bias is to calculate net migration using both the forward and reverse life table survival rate method and averaging the two (Shryock _et al._, 1973, p. 596). The reverse method 'youngs' the population at the time of the later census and compares it to the actual population at the time of the first census. It underestimates mortality and overestimates net migration in emigrating regions by crediting the difference between the projected and actual populations to out-migration when in fact some of those persons died. Thus in emigrating regions, the forward residual method overestimates mortality and underestimates out-migration while the reverse residual method underestimates mortality and overestimates out-migration. In this case the average of the two represents a fairly reasonable approximation of the true value.

This method of averaging the two estimates is inaccurate for regions where in-migration exceeds out-migration. Both forward and reverse residual methods miss those in-migrants who die before they are enumerated. Thus they both underestimate net migration and mortality causing the average of the two to also be an underestimate. This method is perhaps not adequate for calculating net migration in larger Canadian urban areas from 1951 to 1961, most of which experienced in-migration.

A less biased method involves using only the forward migration residual method, which underestimates net migration in both immigrating and emigrating regions, and inflating the migrant population by applying a raising factor equal to the inverse of the life table survival rate for one-half the decade. This method is valid if the assumption can be made that yearly net migration is constant throughout the decade.

The net migration figure is divided by the 1951 labour force size to obtain a measure of the impact of net migration on the urban labour force. This is done because of the likelihood that net migration reflects labour force size. Comparing absolute numbers of migrants is meaningless in cases such as this where city size varies from 25,000 to 2,000,000.

Annexation confounds the validity of this measure as an indicator of in-migration. The size of the 1961 labour force vis-à-vis the 1951 labour force is a function of natural increase, in-migration, and annexation. The three components cannot be separated with available data. It may be argued in favour of the measure that annexation occurs after a considerable number of in-migrants have settled on the periphery of the city. Thus the measure may reflect in-migration over the previous years, which is also a useful indicator to have in the context of this study. However, natural increase also contributes to the measure because residents of the city who enter the work force may also move to the periphery of the city. Thus migration impact is subject to the same weaknesses as residual change in labour force size with respect to annexation.

Average Value Added

An essential property of a growth centre is a rate of economic growth which is more rapid than in other regions. Firms within the centre are able to parlay profits to a greater extent than firms outside the centre for reasons of agglomeration, scale economies or entrepreneurial activity. An additional variable which can be selected to represent growth pole theory is one which reflects 'value added' for manufacturing industries in the urban areas. It is computed in 1951 by deducting the cost of materials, fuel, and electricity consumed from the gross value of products (Dominion Bureau of Statistics, 1965a, p. 14). In 1961 the same deduction is made from value of production which is "the value of shipments adjusted for changes in the value of inventories of finished goods and goods in process" (Dominion Bureau of Statistics, 1965a, p. 8). The alteration in method of calculation is minor, allowing viable comparison for the two census years.

The total value added in each urban area is not a useful variable because it reflects the number of firms and the labour force size. The variable was dismissed when total value added for manufacturing industries in 1951 was seen to correlate 0.94 with labour force size in 1951. A more appropriate measure is average value added per establishment in each urban area. This has the advantage of reflecting not only the profitability of establishments but also their size, because large firms tend to add more value than smaller firms due to economies of scale.

A major limitation of using average value added is that data are only available for manufacturing industries. Value added is not published for other industries by urban area. This indicator may not be entirely reliable in that it may not weight growth centres with non-manufacturing propulsive industries equally with those whose propulsive industries are manufacturing industries. However, most industries in a growth centre are expected to show growth.

The propulsive industry is supposed to produce external economies which contribute to the development of other economic activities in the growth centre (Hansen, 1967, p. 718).

The variable, average value added, is consequently considered a viable indicator of growth centres although it may be biased to some extent towards those urban areas whose propulsive industries are manufacturing industries. The fact that only part of the total industrial structure, manufacturing, is being tapped should be kept in mind while the variable is being analyzed.

Another limitation is that data on value added are not published for those cities in which 75 percent or more of total production is accounted for by one firm or in which 90 percent or more of total production is accounted for by two firms. Consequently, no data on value added are available for five of the urban areas. The five comprise Chicoutimi, Oshawa, Sault Ste. Marie, Sudbury, and Sydney. This reduces the number of units of analysis from forty-six to forty-one with respect to this variable. Data also are not available for Metropolitan Areas of Calgary, Edmonton, Ottawa, and Quebec so Urban Area data are used in place of Metropolitan Area data in these four cases.

Average value added is collected for 1951 and 1961. The 1961 values are regressed on 1951 values to obtain residual change values for each of the forty-one areas. The data are plotted to determine whether their distributions meet the requirements of linear regression analysis. The plot of the 1951 and 1961 values reveals linearity. Plots of the residuals with the independent as well as with observed and expected values of the dependent variable satisfy the requirement for independence.

Further Analysis

The new data set laid out in Table 15 comprises ten original variables plus three new ones representing growth pole theory: 1951 average value added, residual change in average value added, and migration impact on the labour force. Migration impact on the labour force correlates most highly with the factor score summarizing residual change in labour force size and number of retail, wholesale, and service units (0.56), with 1951 average wage of male wage earners (0.52), and with 1961 average wage of male wage earners (0.44). Average value added in 1951 correlates most highly with the factor score summarizing the percentage of the 1951 labour force engaged in manufacturing and in service industries (-0.69), with residual change in location quotient (0.47), and with percentage of the 1951 manufacturing labour force engaged in capital-goods industries (0.43). Residual change in average value added correlates most highly with

Table 15

Simple Correlation Matrix of Fifteen Variables Used in Regression Equations[*]

	Variable[**]													
	1	2	3	4	5	6	7	8	9	10	11	12	13	14
13	0.44	0.52	-0.15	0.06	0.00	0.01	0.14	-0.33	0.27	0.13	0.56	0.14		
14	0.12	0.33	0.14	0.43	-0.30	0.47	-0.00	-0.01	0.03	-0.69	0.00	-0.20	0.12	
15	0.30	-0.01	0.02	-0.11	0.57	0.09	0.09	0.05	0.21	0.04	0.01	0.02	-0.22	-0.00

[*] The remaining correlations are in Table 3

[**]
1. 1961 average wage
2. 1951 average wage
3. 1951 location quotient
4. 1951 % capital goods
5. Change in average wage
6. Change in location quotient
7. Change in % manufacturing
8. Change in % service
9. Change in % capital goods
10. 1951 % manufacturing and service
11. Change in labour force and units
12. 1951 labour force and units
13. Migration impact
14. 1951 average value added
15. Change in average value added

Table 16

Regression Coefficients of Seven Variables and Contributions at Each Step to Reduction of SS of 1961 Wage Variance in the Stepwise Regression

Step	Intercept	Variable* 13	15	4	12	7	11	9	Prop. Variance Reduced	Cmltv. Prop. Variance Reduced
1	3600	5.62							0.198	0.198
2	3563	6.79	0.72						0.163	0.361
3	3363	6.64	0.78	6.75					0.100	0.461
4	3378	6.06	0.75	6.72	111.2				0.093	0.554
5	3372	6.57	0.80	6.38	111.1	-35.16			0.054	0.608
6	3348	7.27	0.83	6.37	111.4	-36.07	-28.05		0.006	0.614
7	3347	7.62	0.86	5.95	110.9	-34.23	-29.67	-3.15	0.004	0.618

* 13. Migration impact
15. Change in average value added
4. 1951 % capital goods
12. 1951 labour force and units
7. Change in % manufacturing
11. Change in labour force and units
9. Change in % capital goods

Table 17

Standardized Regression Coefficients of Seven Variables at each Step of the Stepwise Regression Analysis

	Variable*						
Step	13	15	4	12	7	11	9
1	0.44						
2	0.54	0.41					
3	0.53	0.45	0.32				
4	0.48	0.43	0.32	0.31			
5	0.52	0.46	0.30	0.31	-0.24		
6	0.57	0.48	0.30	0.31	-0.24	-0.09	
7	0.60	0.49	0.28	0.31	-0.23	-0.10	-0.08

*13. Migration impact
15. Change in average value added
4. 1951 % capital goods
12. 1951 labour force and units
7. Change in % manufacturing
11. Change in labour force and units
9. Change in % capital goods

residual change of average wage of male wage earners (0.57), with average wage of 1961 male wage earners (0.30), and with migration impact on the labour force (-0.22).

The average wage of 1961 male wage earners is regressed on all ten variables representing the three theories of urban economic growth. The ten variables are equivalent to those listed in Table 15 minus 1951 average wage of male wage earners, 1961 average wage of male wage earners, and residual change of average wage of male wage earners. The regression coefficients at each step in the stepwise regression procedure for which proportionate reduction in variation of the dependant variable occurs are displayed in Table 16. The corresponding standardized regression coefficients are found in Table 17. Very little reduction in unexplained variance occurs after the fifth step. The best predictors in order of their inclusion in the stepwise regression equation are:

1. migration impact on the labour force (X_{13});
2. residual change in average value added (X_{15});
3. percentage of the 1951 manufacturing labour force engaged in capital-goods industries (X_4);
4. factor score summarizing 1951 labour force size and number of retail, wholesale, and service units in 1951 (X_{12});
5. residual change in percentage of the labour force engaged in manufacturing industries (X_7).

When these five variables comprise the predictors in the regression equation, it takes the following unstandardized and standardized forms:

$$Y = 3372 + 6.57X_{13} + 0.80X_{15} + 6.38X_{4} + 111.1X_{12} - 35.16X_{7}$$

$$Y = 0.52X_{13} + 0.46X_{15} + 0.30X_{4} + 0.31X_{12} - 0.24X_{7}.$$

The first, second, and fourth variables represent growth pole theory while the third and fifth represent sector theory. Once again, the fifth variable, residual change in percentage of the labour force engaged in manufacturing industries, is negative, suggesting that although it is selected to represent sector theory, its negative coefficient negates the theory rather than supporting it. The negative coefficient indicates that the interplay between wages and distribution of labour force in industry categories is more complex than sector theory portrays. It may be that at a certain stage of a region's economic development, increase in proportion of the labour force engaged in manufacturing industries plays a major role in development while at another stage it does not. An intervening factor may be automation which may add to a region's exports while reducing the proportion of its labour force in manufacturing. Another possible factor is a time lag that may occur between increase in manufacturing labour force size and improvement in wage levels. These possibilities are not specified in sector theory.

Table 18

Measures of Importance of Five Variables from Each of Growth Pole Theory and Sector Theory

	β	Proportionate Reduction in Variance	Cumulative Proportionate Reduction in Variance
Sector Theory			
1951 % capital goods	0.40	0.089	0.089
Change in % capital goods	0.18	0.019	0.108
Change in % manufacturing	-0.28	0.024	0.132
Change in % service	-0.16	0.012	0.144
1951 % manufacturing and service	0.12	0.012	0.156
Growth Pole Theory			
Migration impact	0.50	0.198	0.198
Change in average value added	0.40	0.163	0.361
1951 labour force and units	0.34	0.094	0.455
1951 average value added	0.13	0.017	0.472
Change in labour force and units	-0.05	0.002	0.474

Table 18 allows a comparison to be made of the total effectiveness of growth pole theory and sector theory in reducing unexplained variance in the dependent variable. The dependent variable is regressed separately on five variables from each theory. Growth pole theory appears much more powerful than sector theory. It reduces 47.4 percent of the unexplained variance in 1961 average wage of male wage earners while sector theory only reduces 15.6 percent of the unexplained variance. The best single predictor is migration impact on the labour force, while the second best predictor is residual change in average value added. Each of these two variables from growth pole theory reduces more unexplained variance than all five sector theory variables combined.

Migration impact on the labour force is a demographic variable. It is the percentage change in labour force size due to net migration. Demographers might treat this as a dependent variable and attempt to discern which conditions in 1951 and which changes between 1951 and 1961 cause variation in migration impact in that decade.

Migration impact on the labour force is regressed on all 1951 variables and residual change variables displayed in Table 15 to determine the relative powers of these independent variables in reducing unexplained variance in the dependent variable. The most powerful variable is the factor score summarizing residual change in labour force size and number of wholesale, retail, and service units. However, this variable is not conceptually distinct from migration impact on the labour force, so it was dropped from the regression equation and the stepwise regression procedure was repeated. The results are found in Tables 19 and 20.

Inclusion of the first six variables in the stepwise regression procedure accounts for 55.6 percent of all unexplained variation in the dependent variable while inclusion of additional variables accounts for only another 2.3 percent. Thus, taking the first six variables as the predictor variables in the regression equation, we obtain the following unstandardized and standardized regression equations:

$$Y = -155 + 0.08X_2 - 3.21X_8 + 12.68X_{10} - 0.04X_{15} + 0.60X_9 + 0.02X_{14}$$

$$Y = 0.55X_2 - 0.25X_8 + 0.39X_{10} - 0.26X_{15} + 0.19X_9 + 0.21X_{14}$$

where:

Y = migration impact on the labour force;

X_2 = 1951 average wage of male wage earners;

X_8 = residual change in percentage of the labour force engaged in service industries;

Table 19

Regression Coefficients of Eight Variables and Contributions at Each Step to Reduction of SS of Migration Impact Variance in Stepwise Regression

		Variable*								Prop. Variance Reduced	Cumltv. Prop. Variance Reduced
Step	Intercept	2	8	10	15	9	14	3	4		
1	-134	0.07								0.268	0.268
2	-138	0.08	-4.46							0.122	0.390
3	-162	0.09	-4.13	8.48						0.062	0.452
4	-162	0.09	-3.97	8.79	-0.03					0.044	0.496
5	-155	0.08	-3.41	7.83	-0.04	0.68				0.040	0.536
6	-155	0.08	-3.21	12.68	-0.04	0.60	0.02			0.020	0.556
7	-148	0.08	-3.28	10.81	-0.04	0.65	0.02	-1.26		0.014	0.570
8	-156	0.09	-3.13	10.55	-0.04	0.56	0.02	-1.35	-0.20	0.009	0.579

*2. 1951 average wage
8. Change in % service
10. 1951 % manufacturing and service
15. Change in average value added
9. Change in % capital goods
14. 1951 average value added
3. 1951 location quotient
4. 1951 % capital goods

Table 20

Standardized Regression Coefficients of Eight Variables at Each Step of Stepwise Regression Analysis

Step	Variables 2	8	10	15	9	14	3	4
1	0.52							
2	0.53	-0.35						
3	0.60	-0.32	0.26					
4	0.60	-0.31	0.27	-0.21				
5	0.58	-0.27	0.24	-0.25	0.21			
6	0.55	-0.25	0.39	-0.26	0.19	0.21		
7	0.55	-0.26	0.33	-0.26	0.20	0.19	-0.13	
8	0.60	-0.25	0.32	-0.26	0.17	0.22	-0.14	-0.12

2. 1951 average wage
8. Change in % service
10. 1951 % manufacturing and service
15. Change in average value added
9. Change in % capital goods
14. 1951 value added
3. 1951 location quotient
4. 1951 % capital goods

X_{10} = factor score summarizing percentage of the 1951 labour force engaged in manufacturing and service industries;

X_{15} = residual change in average value added;

X_9 = residual change in percentage of the manufacturing labour force engaged in capital-goods industries;

X_{14} = average value added in manufacturing industries in 1951.

Thus we see that migrants are attracted to those areas with high wages. This does not come as a surprise because it is supposed to be one of the functions of regional wage differentials to redistribute the labour force. The third best predictor is variable ten, the factor score summarizing percentage of the 1951 labour force engaged in manufacturing and service industries. This suggests that labour was attracted to those areas which had well developed service and manufacturing sectors. It is possible that these were the areas which were developing and which consequently had job opportunities for migrants.

Summary

This chapter deals with the analysis of data collected to represent variables indicating urban economic development, growth pole theory, sector theory, and export base theory. A number of multiple regression equations were used to determine the relative strengths of predictor variables in accounting for unexplained variation in the dependent variables, which are 1961 average wage of male wage earners and residual change in average wage of male wage earners.

The results of the regression analysis indicate that the two variables central to export base theory, 1951 location quotient coefficient and residual change in the location quotient coefficient, contribute very little to reduction of the unexplained variance in the dependent variables.

The five indicators of sector theory account for more unexplained variance in the dependent variables than did the two from growth pole theory but the pair of growth pole theory variables are better predictors than any pair from sector theory. Three additional variables are drawn from growth pole theory and the two sets of five variables are compared.

Four regression models are finally used. The first model includes five variables from each of growth pole theory and sector theory and two from export base theory. After the model is pared to five predictors, the regression equation takes the form:

$$Y = 3372 + 6.57X_{13} + 0.80X_{15} + 6.38X_4 + 111.1X_{12} - 35.16X_7$$

where:

Y = average wage of 1961 wage earners;

X_{13} = migration impact on the labour force;

X_{15} = residual change in average value added;

X_4 = percentage of the 1951 manufacturing labour force engaged in capital-goods industries;

X_{12} = factor score summarizing labour force size and number of retail, wholesale, and service units in 1951;

X_7 = residual change in percentage of the labour force engaged in manufacturing industries.

These five variables account for 60.8 percent of the total unexplained variance in the dependent variable. When the regression coefficients are standardized, the equation takes the form:

$$Y = 0.52X_{13} + 0.46X_{15} + 0.30X_4 + 0.31X_{12} - 0.24X_7$$

The first, second, and fourth variables represent growth pole theory while the third and fifth represent sector theory.

The second and third regression models allow comparison of the exclusive predictive powers of sector theory and growth pole theory. The five variables from growth pole theory account for 47.4 percent of unexplained variance of the dependent variable, whereas the five variables from sector theory account for only 15.6 percent.

Finally, a regression model is employed which treats migration impact on the labour force as the dependent variable. Paring the model to six independent variables, the regression equation takes the following unstandardized and standardized forms:

$$Y = -155 + 0.08X_2 - 3.21X_8 + 12.68X_{10} - 0.04X_{15} + 0.60X_9 + 0.02X_{14}$$

$$Y = 0.55X_2 - 0.25X_8 + 0.39X_{10} - 0.26X_{15} + 0.19X_9 + 0.21X_{14}$$

where:

Y = migration impact on the labour force;
X_2 = 1951 average wage of male wage earners;
X_8 = residual change in percentage of the labour force engaged in service industries;
X_{10} = factor score summarizing percentage of the 1951 labour force engaged in manufacturing and service industries;
X_{15} = residual change in average value added;
X_9 = residual change in percentage of the manufacturing labour force engaged in capital-goods industries;
X_{14} = average value added in manufacturing industries in 1951.

These six variables account for 55.6 percent of all unexplained variance in migration impact on the labour force. Thus we find that wages and proportions of the labour force in manufacturing and service industries are the best predictors of proportionate influx of migrants into the labour force.

The original concern in this study was with the relative usefulness of growth pole theory, sector theory, and export base theory in explaining urban economic development in Canada from 1951 to 1961. Growth pole theory appears to rank highest with sector theory second and export base theory third. The variables in growth pole theory that appear to be most powerful are indicators of agglomeration and propulsive growth. The most powerful variable from sector theory is the proportion of the manufacturing labour force engaged in capital-goods industries.

Chapter 5

Discussion and Conclusions

The subject of this chapter departs from that of previous chapters which concerned themselves with the merits and demerits of the three theories of regional development. The first part of this chapter deals with the most influential variables used in the study. An elaboration of the relationship between each of the dependent and independent variables is followed by a survey of studies done in Canada to determine what the major factors of urban growth are. The chapter concludes with a suggested framework for a theory of differential urban growth in Canada based on the results of this and other studies and on the recent ideas of urban economists.

Discussion of Major Variables

The variables discussed in this section include those which have correlation coefficients above 0.25 with any of the three dependent variables. These three include the average wage of male wage earners in 1961, residual change in average wage of male wage earners, and migration impact on the labour force. The intercorrelations of the three variables are as follows:

	1	2
1. 1961 average wage		
2. Change in average wage	-0.02	
3. Migration impact	0.44	0.00

Note that the average wage of male wage earners in 1961 and migration impact on the labour force are correlated but that the residual change in average wage of male wage earners does not correlate with the other two.

It is interesting to attempt to relate the three conceptually. If we consider that wages in a given city are a function of the supply and demand for labour, then wages should decline in cases of in-migration and increase in cases of out-migration. Thus we would expect a negative correlation between the residual change in average wage of male wage earners and migration impact on the labour force, assuming a static situation of supply and demand. Yet their correlation coefficient is 0.00. This implies that there is no relationship between wages and migration of workers, an unlikely situation. An explanation of this correlation coefficient is proposed later.

The low correlation between the average wage of male wage earners in 1961 and the residual change in average wage of male wage earners implies that high-wage cities exhibit residual decline in average wage in the preceding decade as often as they show an increase. Similarly, in the set of low-wage cities, as many show an increase in average wage as show a decrease. Thus in the total urban system in 1961 wage differentials tend to be similar to those of 1951 with residual change in the decade occurring randomly with respect to position on the wage hierarchy.

The question arises why migration impact on the labour force is positively correlated with 1961 wages. Migration impact is a reflection of the movement of workers during the decade. It has a correlation coefficient of 0.52 with 1951 average wages and 0.44 with 1961 average wages. Thus those areas with high wages tend to have high in-migration while those with low wages tend to have either low in-migration or some out-migration. This occurs without respect to the size of the urban place. Migration impact on the labour force, which is the proportionate increase in the 1951 labour force due to migration, has a correlation coefficient of 0.14 with the factor score summarizing labour force size and number of retail, wholesale, and service units in 1951. Thus migrants in larger urban areas tend to have a slightly higher proportionate impact on the labour force size than do migrants in smaller areas, although the differential impact is small.

The fact that migration impact on the labour force correlates positively with both size and wage level suggests that size and wage level should in turn be positively correlated. Table 3 reveals that this is so. The factor score summarizing labour force size and number of retail, wholesale, and service units in 1951 has a correlation coefficient of 0.32 with 1961 average wage and 0.17 with 1951 average wage.

These correlations tend to support the following model. Cities with higher wages tend to maintain their position in the hierarchy of wage levels of urban areas. Furthermore, the wage differentials would tend to increase were it not for differential migration rates. In-migration acts to pull down wages in high-wage cities which would otherwise rise even higher while out-migration from low-wage areas has the opposite effect of hindering further decline of wages in low-wage areas. Migration of labour acts to restrict the range of wages in the urban wage hierarchy which would otherwise expand.

This also suggests that increase in size is a reliable indicator of economic development despite Richardson's cautionary statement that population increase is not equivalent to urban development (Richardson, 1971, p. 80). Thompson argues that if we take the long run view, a city does not get richer, it gets bigger (Thompson, 1968, p. 52). Thus increase in size, or migration impact, may be the more viable measure of growth.

The set of three dependent variables provides indication of two independent phenomena. The first phenomenon, the persistence of the hierarchy of urban places according to wage, is supported by the correlation coefficient of 0.85 between the average wage of male wage earners in 1951 and the average wage of male wage earners in 1961 (Table 3). Other variables, such as migration impact on the labour force and the factor score summarizing labour force size and number of retail, wholesale, and service units in 1951, tend to correlate positively with both 1951 and 1961 average wage, indicating that they exact an equilibrating force on the wage hierarchy system.

The third dependent variable, the residual change in average wage of male wage earners, reflects the second phenomenon which is a set of minor variations in wage levels within the hierarchy - variations which are uncorrelated with the wage hierarchy. This may explain why migration impact on the labour force and the residual change in average wage of male wage earners are not correlated. It is because the manifestation of the interplay of migration and wages occur in the correlation between the wage hierarchy and migration rather than between changes in the wage hierarchy and migration. Cities with higher wage levels have accrued a number of economic advantages over other cities and these advantages tend to interact to provide more advantages. The factor that tends to equalize wages is migration.

Table 21

Average Earnings of Male Wage Earners, 1951 and 1961, and Change in Rank, for Canadian Cities

1951			1961			Change in Rank
1.	Welland	$2639	1.	Sarnia	$4662	3- 1
2.	Oshawa	2637	2.	Sault Ste. Marie	4504	4- 2
3.	Sarnia	2630	3.	Ottawa	4407	11- 3
4.	Sault Ste. Marie	2602	4.	Sudbury	4330	6- 4
5.	St. Catherines	2565	5.	Toronto	4330	8- 5
6.	Sudbury	2560	6.	Calgary	4256	14- 6
7.	Windsor	2556	7.	Hamilton	4251	9- 7
8.	Toronto	2529	8.	Oshawa	4248	2- 8
9.	Hamilton	2498	9.	Vancouver	4219	16- 9
10.	Peterborough	2450	10.	St. Catherines	4130	5-10
11.	Ottawa	2416	11.	Peterborough	4099	10-11
12.	London	2413	12.	Belleville	4068	15-12
13.	Chatham	2401	13.	Edmonton	4059	18-13
14.	Calgary	2374	14.	Regina	4012	25-14
15.	Belleville	2358	15.	Windsor	4002	7-15
16.	Vancouver	2352	16.	London	4000	12-16
17.	Kitchener	2345	17.	Montreal	3972	26-17
18.	Edmonton	2340	18.	Kitchener	3950	17-18
19.	Moncton	2333	19.	Kingston	3938	29-19
20.	Port Arthur	2329	20.	Welland	3925	1-20
21.	Brantford	2312	21.	Winnipeg	3907	23-21
22.	Timmins	2287	22.	Saskatoon	3896	33-22
23.	Winnipeg	2278	23.	Port Arthur	3888	20-23

Table 21 (concluded)

1951			1961			Change in Rank
24.	Cornwall	$2276	24.	Chicoutimi	$3883	32-24
25.	Regina	2274	25.	Guelph	3737	31-35
26.	Montreal	2271	26.	Chatham	3729	13-26
27.	Lethbridge	2257	27.	Victoria	3698	34-27
28.	Sydney	2242	28.	Moose Jaw	3688	30-28
29.	Kingston	2234	29.	Halifax	3679	36-29
30.	Moose Jaw	2232	30.	Moncton	3678	19-30
31.	Guelph	2226	31.	Brantford	3667	21-31
32.	Chicoutimi	2219	32.	Lethbridge	3666	27-32
33.	Saskatoon	2214	33.	Cornwall	3646	24-33
34.	Victoria	2203	34.	Shawinigan	3616	37-34
35.	Brandon	2171	35.	Brandon	3594	35-35
36.	Halifax	2167	36.	Quebec	3559	43-36
37.	Shawinigan	2078	37.	St. John's	3515	41-37
38.	Drummondville	2059	38.	Timmins	3450	22-38
39.	Trois Rivières	2034	39.	Trois Rivières	3448	39-39
40.	Sherbrooke	1973	40.	Sydney	3392	28-40
41.	St. John's	1942	41.	St. Jean	3226	44-41
42.	Valleyfield	1935	42.	Sherbrooke	3176	40-42
43.	Quebec	1909	43.	Granby	3157	45-43
44.	St. Jean	1896	44.	Valleyfield	3135	42-44
45.	Granby	1872	45.	Saint John	3110	46-45
46.	Saint John	1846	46.	Drummondville	3030	38-46

Thus migration tends to hold constant an otherwise increasing wage differential, resulting in the positive correlation occurring between migration and wage levels rather than between migration and wage changes.

In Table 21 most cities maintain a relatively stable status in the wage hierarchy from 1951 to 1961. Only seven changed their rank by more than ten. Saskatoon and Regina are the only ones to show such a sizable increase while three cities from Ontario, Welland, Chatham, and Timmins, and two from the Maritimes, Moncton and Sydney, experience decreases of this magnitude. The purpose of this table is to provide the reader with a sense of the consistency of the urban hierarchy over the decade. It is not designed for analytical purposes. The shifts in rank within the hierarchy is not a valid measure of change. It is an ordinal measure subject to regression effects, so further analysis of change will be dealt with in connection with Table 24 which shows residual change in wages, rather than shifts in rank as presented here.

Table 22 presents the three dependent variables used in this study along with those independent variables which have correlation coefficients of more than 0.25 with each of the dependent variables. The intercorrelations of the dependent variables, discussed previously, suggest that the residual change in average wage of male wage earners is

Table 22

Correlation Coefficients of Major Independent Variables With Three Dependent Variables

Dependent Variable	Independent Variable	Correlation Coefficient
Residual change in average wage	1. 1951 % manufacturing and service	0.34
	2. 1951 labour force and units	0.33
	3. 1951 average value added	-0.30
	4. Change in average value added	0.57
Migration impact	1. 1951 average wage	0.52
	2. Change in % service	-0.33
	3. Change in % capital goods	0.27
	4. Change in labour force and units	0.56
1961 average wage	1. 1951 % capital goods	0.35
	2. Change in location quotient	0.30
	3. Change in labour force and units	0.28
	4. 1951 labour force and units	0.32
	5. Change in average value added	0.30
	6. Migration impact	0.44

independent of the other two which are correlated. Thus a model can be constructed with two dimensions: one dimension is represented by migration impact on the labour force and the average wage of male wage earners in 1961 while the other is represented by the residual change in average wage of male wage earners.

The model presented here implies that there are two distinct sets of causes of the two separate dimensions. Independent variables associated with the average wage of male wage earners in 1961 and migration impact on the labour force are considered to be factors that maintain the urban hierarchy of wage levels. Those independent variables associated with the residual change in average wage of male wage earners are considered factors of either independent variation in urban development or of a secular reordering of the urban hierarchy of wage levels. A time series study involving more than two points in time is necessary to distinguish between secular shifts and individual variations in the urban structure. The difference between the two processes lies in their permanency. A secular shift is a long term trend while individual variations are temporary fluctuations which do not disrupt the long term pattern.

The division of dependent variables into equilibrating and disequilibrating effects is not completely justified by the distribution of their correlation coefficients with independent variables. If the two effects are distinct, they should be the result of two distinct sets of forces, those that contribute to stability and those that contribute to instability in the system. But the residual change in average wage of male wage earners shares two independent variables with the average wage of male wage earners in 1961, viz., residual change in average value added and the factor score summarizing labour force size and number of retail, wholesale, and service units in 1951, which suggests that perhaps the division is not completely valid. However, the values of the correlation coefficients are low enough, about 0.33 in three of the four cases, that they do not invalidate the suggested model. A correlation coefficient of 0.33 can be interpreted as meaning that the dependency exists in 0.33^2 or in one out of nine cases. Possibly intervening variables can be invoked to explain the conditions under which the effects of the independent variables are reversed.

At this point it is opportune to point out that few of the correlation coefficients in Table 22 are large enough to establish strong relationships between the dependent and independent variables. This renders the discussion in this section somewhat speculative. It should be kept in mind that a large number of various factors apparently influence wages and that the variables discussed in this study are a small part of a larger system of factors.

It would also be useful to pause and discuss the accepted economic theory of wages since wages are the primary dependent variables in this study; migration impact on the labour force is considered closely connected with wages.

The prominent theory of wages in economics, called marginal productivity theory, has the forces of competition tending to make employers pay labour a wage approximately equal to the full value of its marginal contribution to the revenue of the firm (Sultan, 1957, pp. 475-480). If the marginal product of an additional worker, which is the increase in production incurred because of the presence of that worker, is greater than his wage, the employer is likely to hire another worker in order to realize higher profits. Alternatively, if the additional worker's marginal product lies below his wage, the employer is impelled to increase marginal productivity, reduce wages, or not hire the additional worker.

Marginal productivity theory does not apply with absolute precision, but rather it describes a tendency for wages to adjust themselves according to the margin of productivity. Its critics point out that profit-maximization is not the sole force operating on employer decisions. They also argue that the theory ignores a number of other factors of wage levels such as imperfections of competition, discontinuities of production and cost functions, and the existence of ranges within which bargaining power operates

(Sultan, 1957, p. 482).

The model used in this study is subject to the same criticisms directed at marginal productivity theory and is easily reconciled with that theory. Both make similar assumptions. They assume free competition for labour, knowledge of marginal productivity, and mobility of capital and labour. Both make wages a function of productivity. Productivity is a function of innovation, entrepreneurial activity, efficiency, capital investment, and technological advancement. The independent variables in the model are indicators of productivity according to each of the three theories of regional development just as wages are treated as a function of productivity in marginal productivity theory.

In Table 22, migration impact on the labour force and the average wage of male wage earners in 1961 are treated as two correlated variables dependent on forces of equilibrium, forces that act to maintain the wage hierarchy of Canadian cities. Migration impact on the labour force is considered both as a dependent variable and as an intervening variable acting upon the average wage of male wage earners in 1961. Thus independent variables which correlate with migration impact on the labour force shall be considered to be acting through that variable on the average wage of male wage earners in 1961.

The largest correlate of migration impact on the labour force is the factor score summarizing residual change in labour force size and in the number of retail, wholesale, and service units. This association is not insightful because both variables represent dynamic aspects of the agglomeration process during the decade.

The second highest correlate with migration impact on the labour force, the average wage of male wage earners in 1951, reflects a basic tenet of any theory of migration, viz., migration is influenced by economic opportunities. People migrate from places with job shortages to places with job surplusses. Excess jobs create higher wages according to the law of supply and demand, resulting in a positive correlation between wages and in-migration.

The third largest correlate of migration impact on the labour force is residual change in percentage of the labour force engaged in service industries. It is a negative correlation, indicating that higher in-migration tends to occur in those places with negative residual changes in the service industry.

This negative relationship may be a function of the particular stage of development Canadian cities were undergoing in 1951-1961. The period 1951 to 1961 catches cities at different stages of industrialization and post-industrialization. Perhaps those cities which had already achieved development of their service sector, and consequently were no longer expanding in that sector, were experiencing in-migration because of their advanced development. Other urban areas were in the process of developing their service sectors during the fifties and may show more growth in the following decade. One might argue that the low correlation between percentage of the labour force engaged in service industries in 1951 and each of the indicators of development eliminates this possibility. However, if we can assume that each urban area has a unique optimum size of service sector, then the low correlation between percentage of the labour force engaged in service industries in 1951 and urban development does not conflict with the hypothesis that cities had achieved different levels of development in their service sectors in the Fifties.

A complementary explanation looks to the Korean War, which broke out in 1950, as a causal agent. The Korean War stimulated Canada to invest heavy capital expenditures in resource industries to meet external demand (Royal Commission, 1957, p. 91). Heavy investments were made in the iron ore industry in Labrador and in the development of oil and natural gas on the prairies. A new large-scale aluminum project was undertaken at Kitimat in northern British Columbia. Between 1949 and 1955 output in the resource industries increased by approximately 70 percent (Royal

Commission, 1957, p. 91). This diversion of capital into resource industries in the early fifties may cause the negative relationship between migration and residual change in percentage of the labour force engaged in service industries.

Another possible explanation is that provincial and federal governments selectively expanded military and administrative units into lesser developed areas to stimulate growth or at least to counteract stagnation. If such methods are not immediately effective, and there is evidence that they are not, the result might be a negative correlation between these two variables.

These are speculations. It seems incongruous for residual change in percentage of the labour force engaged in service industries to be negatively correlated with migration impact on the labour force when migration impact on the labour force is positively correlated with the average wage of male wage earners in 1961. The negative relationship is more perplexing because it occurs during the 1950's when the service sector of Canada increased from 44 percent to 55 percent of the total labour force (Dominion Bureau of Statistics, 1953, Table 16; Dominion Bureau of Statistics, 1963b, Table 1A).

The remaining variable that is correlated with migration impact on the labour force is residual change in the percentage of the manufacturing labour force engaged in capital-goods industries. This variable also correlates 0.24 with residual change in percentage of the labour force engaged in manufacturing industries, suggesting that expansion in capital-goods industries played a role in the development of the manufacturing sector. Areas with changing proportions of their labour force in manufacturing tend to show similar proportionate changes in capital-goods industries. Migration impact on the labour force correlates 0.27 with residual change in the percentage of the manufacturing labour force engaged in capital-goods industries while it correlates 0.14 with residual change in percentage of the labour force engaged in manufacturing industries. Thus an increase in percentage manufacturing is associated with in-migration but the association is considerably stronger when that increase occurs in the capital-goods sector of manufacturing.

Hoffman's hypothesis is the only one from sector theory borne out by Canadian data. He states that capital-goods industries require a more elaborate production process than consumer-goods industries. Elaboration results in specialization, more linkages, and a higher input in capital, resulting in greater output per unit input.

This process is supposed to occur at a specific stage of growth according to sector theory. It occurs during the later stages of industrialization, after consumer-goods industries have become established and before the final stage of specialization in services. This may help to explain the negative correlation between residual change in percentage of the labour force engaged in service industries and migration impact on the labour force. Perhaps those cities which expanded their service sectors during the fifties did so prematurely and suffered a comparative loss in labour force.

The association of residual change in the percentage of the manufacturing labour force engaged in capital-goods industries and migration impact on the labour force may reflect the stage of development of the Canadian urban economy. It need not be construed as proof that sector theory is inadequate in describing Canada's urban economic development. Perhaps Canada had attained the final stage of industrialization and was preparing to advance into post-industrial growth. If we accept this position, we face the difficult task of reconciling the negative correlation coefficient of migration impact on the labour force and residual change in percentage of the labour force engaged in service industries. The coefficient should be at least modestly positive, unless it can be argued that there is a time lag between initial development of the service sector and urban development.

This hypothesis conflicts with data for the total Canadian economy for that period. The percentage of the manufacturing labour force in capital-goods industries decreased from 33 percent in 1951 to 30 percent in 1961

(Dominion Bureau of Statistics, 1953, Table 16; Dominion Bureau of Statistics, 1963b, Table 1A). Thus the total economy was not increasing its capital-goods sector; it was increasing its service sector. The problem remains why those cities with the most rapidly expanding service sectors tend to experience little migration impact on the labour force while those cities with an expanding capital-goods sector tend to experience more in-migration. Of course, correlations at the national level do not imply similar correlations at the urban level. Such a statement is subject to the ecological fallacy.

Proportionately high in-migration occurs in cities with higher wages in 1951 and in cities whose manufacturing sectors shift towards greater emphasis on capital-goods industries. This cluster of variables all contribute to the average wage of male wage earners in 1961. The correlation coefficient of the average wage of male wage earners in 1951 and the average wage of male wage earners in 1961 is 0.85, indicating a fairly stable wage hierarchy among the cities of Canada from 1951 to 1961. However the average wage of male wage earners in 1951 is not considered a direct cause of the average wage of male wage earners in 1961. Rather it is a reflection that forces operating on wages in 1961 were also in effect in 1951. These stable inter-urban wage differentials are apparently maintained by the independent variables associated with 1961 and 1951 wage levels. Higher rates of in-migration occur in cities with higher wages. The laws of supply and demand imply that migration acts to hold the range of the wage hierarchy relatively constant. Without migration, wages in higher wage areas would rise still higher while those in lower wage areas would decline even more.

So far we have treated migration impact on the labour force as a dependent variable. When the average wage of male wage earners in 1961 is taken as the dependent variable, migration impact on the labour force is seen to be the most influential independent variable. The factor score summarizing residual change in labour force size and in the number of retail, wholesale, and service units, which also correlates with the average wage of male wage earners in 1961, represents an alternative to migration impact on the labour force as a measure of agglomeration. The factor score summarizing residual change in labour force size and in the number of retail, wholesale, and service units reflects the effects of both in-migration and natural increase on the size of the labour force. It correlates 0.28 with the average wage of male wage earners in 1961, 0.31 with the average wage of male wage earners in 1951, and 0.56 with migration impact on the labour force.

Exploring other correlates of the average wage of male wage earners in 1961, we find that the second highest correlate is percentage of the manufacturing labour force engaged in capital-goods industries in 1951. This variable has a correlation coefficient of 0.35 with the average wage of male wage earners in 1961 and 0.50 with the average wage of male wage earners in 1951. These coefficients reinforce the high correlation of migration impact on the labour force and residual change in the percentage of the manufacturing labour force engaged in capital-goods industries. Apparently Hoffman's hypothesis does apply to the Canadian economy of 1951-1961. Cities with high wage levels tend to have a larger proportion of their manufacturing labour force in capital-goods industries while cities which expand their capital-goods sector tend to experience in-migration.

The third highest correlate of the average wage of male wage earners in 1961 is the factor score summarizing labour force size and number of retail, wholesale, and service units in 1951. It correlates 0.32 with the average wage of male wage earners in 1961 and 0.17 with the average wage of male wage earners in 1951. This variable reflects the amount of agglomeration that occurred previously to the time of the period being studied while the factor score summarizing residual change in labour force size and in the number of retail, wholesale, and service units reflects the process of agglomeration during the period of study. Apparently both the level and the dynamics of agglomeration contribute to equilibrate the wage hierarchy of the system of Canadian cities. Large cities tend to pay higher wages which attract migrants who contribute to further agglomeration as well as to further economic development of the city. However, we must keep in mind that it is not the migrants per se who contribute to urban development, but the availability of jobs created by expanding industry and by the increase in consumption generated by the migrants. Agglomeration is the massing together of firms for their mutual economic benefit. It is not the massing together of migrants looking for jobs. In-migration is a reflection of an increasing number of job opportunities created by agglomeration.

The fourth ranking correlate with the average wage of male wage earners in 1961 is residual change in the location quotient coefficient. It correlates 0.30 with the average wage of male wage earners in 1961 and 0.34 with the average wage of male wage earners in 1951. The only other representative of export base theory, the 1951 location quotient coefficient, correlates 0.01 with the average wage of male wage earners in 1961 and 0.03 with the average wage of male wage earners in 1951. Apparently the size of the export base in 1951 had no effect on the level of development attained in either 1951 or 1961. But changes in the export base did have an effect in development attained in 1961. A possible explanation for the lack of correlation in one case and the modest positive correlation in the other is that rather than the export base influencing wages, higher wages influenced the export base. Those cities with higher wages also have larger agglomeration and more entrepreneurs providing them with greater economic flexibility. They are able to adjust their economic base more quickly to capitalize on external demand and thus to maintain their status on the wage level hierarchy.

The final contributor to the average wage of male wage earners in 1961 is residual change in average value added with a correlation coefficient of 0.30. Residual change in average value added is selected to tap those growth centres with profitable industries. It seems reasonable to expect that those cities with higher profit firms pay higher wages than those with lower profit firms.

To recapitulate, it is suggested that wage levels are stabilized by means of a number of factors, primarily migration. Factors which contribute to migration and therefore to stable wage levels are profitability, agglomeration, and fluctuations in the export base. Cities which have attained higher economic development tend to maintain this level by means of a number of interacting forces. Agglomeration contributes to profitability and to entrepreneurial activity, allowing the more developed city to expand in capital-goods industries and export base industries in order to maintain its economic superiority.

Within this relatively stable hierarchy there are some variations. These variations are reflected by the residual change in average wage of male wage earners which is a measure of the degree of variation in urban development experienced by individual cities. It derives from residual values obtained when the average wage of male wage earners in 1961 is regressed on the average wage of male wage earners in 1951. This variation may be random, it may reflect temporary minor adjustments, or it may reflect long term secular shifts in the urban hierarchy. The ten year period under study is not long enough to distinguish among these different types of change.

The largest correlate of the residual change in average wage of male wage earners is residual change in average value added. Their correlation coefficient is 0.57, the largest in Table 22. This indicates a relationship between wages and profits. Employees in firms with increasing profits tend to receive an increase in wages while those in firms with decreasing profits may receive declining wages through less overtime pay or smaller bonuses. It is seldom that wage rates actually decline. The causal arrow may have two heads: firms may experience greater profits after increasing wages, or, more likely, after hiring more highly skilled personnel.

Since residual change in average value added is supposed to indicate profitability of firms, and since profitability is one of the goals of firms, it seems plausible to treat residual change in average value added as a dependent rather than an independent variable. This might be useful in other studies but in the present one residual change in average value added does not correlate highly with any variables except the residual change in average wage of male wage earners and the average wage of male wage earners in 1961, both of which are considered dependent variables.

The second highest correlate of the residual change is average wage of male wage earners in the factor score summarizing percentage of the labour force engaged in manufacturing and service industries in 1951 with a correlation coefficient of 0.34. This suggests that those cities with larger positive residual change in average wage of male wage earners tend to have larger proportions of their labour force in manufacturing and service industries, while those cities with larger negative residual change in average wage of male wage earners tend to have smaller proportions of their labour force in these industries.

The third highest correlate of the residual change in average wage of male wage earners is the factor score summarizing labour force size and number of retail, wholesale, and service units in 1951. It has a correlation coefficient of 0.33 with the residual change in average wage of male wage earners and 0.32 with the average wage of male wage earners in 1961. This constitutes a paradox in that agglomeration contributes to both the equilibrium and the disequilibrium of the Canadian urban system. The 'paradox' is not serious when the extent of the association, 0.33 and 0.32, is noted. Some cities may maintain their positions in the wage hierarchy because of advantages or disadvantages of their size while others deviate from their positions in the hierarchy because of their size. Further relationships between size and development are considered when Tables 23 and 24 are discussed.

The final correlate with the residual change in average wage of male wage earners is also perplexing. It is average value added in 1951, which correlates negatively with the residual change in average wage of male wage earners, -0.30. This means that there is a tendency for cities with higher average value added in 1951 to experience a residual decline in wages over the following ten years and/or for cities with lower average value added in 1951 to experience a residual increase in wages over those years.

A possible explanation is that 1951 was a year of rapid inflation generated largely by the Korean War. Inflation can be a disruptive factor in the profits of certain industries. Some may temporarily benefit from rising prices while others are not able to adjust so quickly and may experience a temporary decline in profits. Average value added is particularly sensitive to the national and international economic climate - more so than other, 'sticky' variables such as labour force size, wages, and percentages of the labour force in various industry sectors. Thus the period of inflation in the early Fifties may attenuate some of the relationships between 1951 average value added and the change variables and 1961 variables.

In summary, the variables selected in this study are inadequate for a comprehensive explanation of differential urban development in Canada from 1951 to 1961 because they are not selected for that purpose. Rather, they are selected to represent crucial variables in each of three theories of regional development to determine the extent to which each theory explains urban development. The conclusion is that each theory makes some contribution, and that growth pole theory is superior to sector theory which in turn is superior to export base theory.

This section has aimed at evaluating the individual variables as to their efficiency in explaining the urban structure and changes in the urban structure. It is surmised that the structure, as defined by wage levels, is stabilized largely by migration. The major factors of migration are job opportunities. These in turn are determined by higher wage levels, agglomeration, and development of capital-goods industries. Inter-urban wage differentials are maintained by migration, profitability, agglomeration, and the ability to expand in profitable industries.

Contributions to a Theory of Differential Urban Development in Canada

The selection of variables has so far been necessarily restrictive because of the intention to compare three theories of regional development. Although no single available theory is apparently adequate for the purpose of explaining urban development, a number of studies have been carried out which attempt to grasp underlying causes of urban development in Canada. These are discussed in the following paragraphs.

The first concern is with broad regional differences in Canada and their effects on the city system. One possible reason for a city's rate of growth is its geographic setting. If a region has the resources, capital and locational advantage for development, its cities will also develop.

A number of other factors should be considered in addition to the variables revealed in this study to be useful for the development of a theory of differential urban growth in Canada. Some crucial factors, as indicated by Canadian research, include regionalism, a heartland-hinterland concept, the concept of an economic shadow, and centralization of financial functions within the total economy. These will be discussed in the following sections.

Regionalism

The first concern is with broad regional differences in Canada and their effects on the city system. One possible reason for a city's rate of growth is its geographic setting. An indication that the urban hierarchy can advantageously be viewed in terms of Canadian regions is provided by regional comparisons of personal income per capita from 1926 to 1964. Expressed as a ratio of the Canadian average, first place has been exchanged back and forth by British Columbia and Ontario, averaging 125 percent of the national average (Baxter, 1966, p. 2). Quebec has been fairly constant at 80-90 percent over that period of time as has the Maritimes with 60-70 percent of the national average. Only the prairie provinces have shown considerable fluctuation, from a low of 70 percent in the Thirties to a high of almost 120 percent in the Fifties. These regional disparities are mirrored in a number of other indicators such as standard of the work force as measured by educational level, and annual investment per capita (Baxter, 1966, p. 2).

A crude method of determining whether regions influence urban development is to group cities according to their location in the West, Ontario, Quebec, and the Maritimes, and to compare the distribution of cities in each region with respect to indicators of development.

Table 23

Percentage Change in Labour Force Size Attributed to Migration

1.	Cornwall	102
2.	St. Catherines	93
3.	Welland	93
4.	Calgary	86
5.	Edmonton	81
6.	Saskatoon	65
7.	London	64
8.	Sudbury	63
9.	Drummondville	56
10.	Toronto	53
11.	Kitchener	47
12.	Regina	45
13.	Kingston	42
14.	Vancouver	41
15.	Ottawa	40
16.	Lethbridge	40
17.	Moncton	38
18.	Belleville	37
19.	Brantford	34
20.	Montreal	33
21.	Hamilton	33
22.	Brandon	33
23.	Guelph	32

Table 23 (concluded)

24.	Halifax	29
25.	Port Arthur	28
26.	Oshawa	28
27.	Winnipeg	26
28.	Sarnia	25
29.	Moose Jaw	24
30.	Chatham	24
31.	Sault Ste. Marie	16
32.	Granby	15
33.	Quebec	12
34.	Sherbrooke	5
35.	Victoria	4
36.	Peterborough	3
37.	St. Jean	2
38.	Windsor	1
39.	St. John's	0
40.	Saint John	-5
41.	Chicoutimi	-5
42.	Trois Rivières	-6
43.	Valleyfield	-8
44.	Shawinigan	-11
45.	Timmins	-14
46.	Sydney	-27

Table 24

Residual Change in Average Wage of Male Wage Earners in Canadian Cities, 1951-1961*

1.	Ottawa	$376
2.	Quebec	311
3.	Sarnia	300
4.	Calgary	290
5.	Vancouver	287
6.	St. John's	216
7.	Regina	200
8.	Kingston	188
9.	Sault Ste. Marie	185
10.	Saskatoon	177
11.	Montreal	165
12.	Chicoutimi	156
13.	Edmonton	145
14.	Belleville	126
15.	Toronto	124
16.	Shawinigan	107
17.	Hamilton	93
18.	Winnipeg	89
19.	Sudbury	76
20.	Halifax	32
21.	Kitchener	28
22.	Trois Rivières	22
23.	Peterborough	15

*Obtained by regressing 1961 average wage on 1951 average wage

Table 24 (concluded)

24.	Guelph	$-1
25.	St. Jean	-2
26.	Victoria	-4
27.	Port Arthur	-9
28.	London	-27
29.	Granby	-39
30.	Saint John	-41
31.	Brandon	-59
32.	Moose Jaw	-59
33.	Lethbridge	-120
34.	Oshawa	-125
35.	St. Catherines	-132
36.	Valleyfield	-153
37.	Cornwall	-169
38.	Sherbrooke	-171
39.	Brantford	-204
40.	Moncton	-225
41.	Windsor	-246
42.	Chatham	-279
43.	Sydney	-370
44.	Timmins	-382
45.	Drummondville	-450
46.	Welland	-451

Table 23 ranks the cities according to migration impact on the labour force while Table 24 ranks them according to residual change in average wage of male wage earners during the decade. These two tables along with Table 21, which ranks the cities according to average wage of male wage earners in 1951 and 1961, provide four indicators of urban development.

The indicator of urban development in Table 24 supposedly reflects disequilibrating change, change that disrupts the established wage hierarchy. Comparing the cities in the four regions as to their positive or negative residual change in average wage of male wage earners we obtain the following distributions by region:

	West	Ontario	Quebec	Maritimes
Positive change	6	10	5	2
Negative change	4	11	5	3

Although the West is the only region to show more positive than negative residual change in average wage of male wage earners, the differences between the regions are marginal. On the basis of the indicators chosen, Canadian regions do not appear to exert a disequilibrating force on the urban hierarchy of wages.

Turning to indicators of the established wage hierarchy, we observe that the association of migration impact on the labour force and regional affiliation of the city in Table 23 is as follows:

	West	Ontario	Quebec	Maritimes
Above median	7	13	2	1
Below median	3	8	8	4

Table 23 reveals that there are very significant regional differences in migration impact on the labour force among the cities. Almost two-thirds of western Canadian and Ontario cities lie above the median, i.e., experience more than a thirty percent increase in labour force size due to in-migration. Only twenty percent of Quebec and Maritime cities experience such a high rate of in-migration.

Table 21 contains two sets of indicators of the urban wage hierarchy, average wage of male wage earners in 1951 and average wage of male wage earners in 1961. The distribution of cities in regions by average wage in 1951 is as follows:

	West	Ontario	Quebec	Maritimes
Above median	4	18	0	1
Below median	6	3	10	4

Extreme regional disparities in average male wage exist in 1951. Ontario is the only region with more cities above the median, which is $2277, than below it. Only five cities outside Ontario, Moncton, Edmonton, Calgary, Winnipeg, and Vancouver, have average male wages above the median. The thirteen highest ranking cities, and fifteen of the top seventeen, are located in Ontario. At the other end of the spectrum, the lowest ranking eleven cities are located in Quebec and the Maritimes.

The distribution of cities in regions by average wage of male wage earners in 1961 is a little more egalitarian:

	West	Ontario	Quebec	Maritimes
Above median	6	16	1	0
Below median	4	5	9	5

The median has increased to $3885 but the regional disparities remain very similar to those of 1951. Once again, only one city east of Ontario is above the median, and it is now Montreal, located close to Ontario's borders. As a point of techicality, part of the Metropolitan Area of Ottawa is located east of the Ottawa River but the large part of the Metropolitan Area is located within Ontario and it is considered an Ontario city. The first five cities, and ten of the first twelve, are located in Ontario. Again the eleven lowest ranking cities are located east of Ontario.

Tables 21 and 23 reveal definite regional disparities in the established wage hierarchy of Canadian cities. Ontario ranks highest while western Canada ranks second. On the strength of the performance of Montreal, Quebec ranks third and the Maritimes fourth. The distribution of residual changes of average wage in Table 24 suggests that the hierarchy is not in danger of disruption.

The broad regions of Canada appear to have cultural as well as wage differentials. Ray (1971) factor analyzed 84 variables from the 1961 census for all cities with populations over 10,000. The first, second, and fifth ranking factors have loadings which lead him to name them 'English-French contrasts', 'Maritime-prairie contrasts', and 'British-Columbia-type cities'. The variables with the highest loadings on English-French contrasts are: percentage of the population with English mother tongue (+0.941), percentage of the population of the United Church (+0.918), percentage of the population with French mother tongue (-0.908), and percentage of the population Roman Catholic (-0.898) (Ray, 1971, p. 6).

The variables loading highest on the Maritime-prairie factor are: percentage of the population born in East Europe (-0.751), distance from Vancouver (+0.739), and percentage of the population with East European mother tongue (-0.733) (Ray, 1971, p. 9). The prairie provinces show the most cultural diversity while the Maritimes show the least.

The variables loading highest on British Columbia-type cities are: percentage of the population born in British Columbia (-0.799), percentage of the population born in Scandinavia (-0.785), distance from Toronto (-0.747), and percentage of the population with Asian mother tongue

(-0.714) (Ray, 1971, p. 11).

No occupation-income characteristics are associated with the English-French factor. The Maritime-prairie factor contrasts the proportion of farmers and labourers in the labour force. The British Columbia factor describes almost as many economic as cultural characteristics, including the percentage of fishermen, longshoremen, carpenters, and loggers in the labour force (Ray, 1971, p. 11).

King (1966) uses principal-component analysis and a grouping algorithm to obtain regional groupings of 106 cities that had populations of 10,000 or more in 1951. His groups are investigated here in the hope that they may provide insight into the urban wage structure and its changes. He derives twelve components from 52 variables measuring economic, social, demographic, and locational characteristics. His first component is an index of the youthfulness of the female population in non-manufacturing and less densely populated cities. The second component, which is a socio-economic indicator, appears to index the service role of many comparatively isolated communities located outside Quebec. The third component indexes frontier location and a closer economic orientation to primary activities while the fourth is identified with urban manufacturing in the smaller cities located outside the major metropolitan complexes (King, 1966, p. 209).

The fifth component contains two aspects of urban structure. The first is the occupational and housing correlates of the suburban communities while the second is a facet of manufacturing employment and of the role of the city as a port (King, 1966, p. 213). The sixth component is an employment-housing dimension, reflecting an emphasis upon manufacturing employment in metals production, and comparatively high proportions of occupied dwellings in need of repair (King, 1966, p. 213). King is unable to interpret and label the remaining six components.

King groups the cities according to their location in the twelve-dimensional orthogonal space of the twelve components. He uses a matrix of distances between cities in the space to group those cities that are proximate. He derives eleven groups which he finds to be consistent with the accepted broad regional divisions of the country (King, 1966, p. 214). The number of groups is arbitrary. The method successively reduces the number of groups from 106 to one. King chose the number of groups to be eleven because at that point he has eliminated all but one single member groups.

King's set of cities is larger than the 46 included in the present study. They also differ in that he does not aggregate those cities which are geographically contiguous, as is done here. Nevertheless the focus of his study is relevant to the concerns of this one and his cities include the ones studied here over the same period of time. His sophisticated statistical analysis deserves considerable scrutiny to determine whether it contributes to an understanding of causes and consequences of the urban wage hierarchy and its fluctuations.

The 46 cities in this study are placed in seven of the eleven groups derived by King. Table 25 substantially supports King's contention that the groupings coincide with broad regional divisions. Groups 1 and 6 contain cities from Quebec while all of group 3 cities are in southern Ontario. Group 4 cities lie west of the Great Lakes. Groups 2 and 5 are multi-regional. Group 2 are all resource industry cities while group 5 cities tend to be older and/or larger cities. The remainder of the groups will be ignored in this analysis. This eliminates Kitchener from the analysis. It is the only city that does not fall in one of the six groups mentioned above.

An important question to ask is why the statistical methods used by King cause the cities to be grouped this way. The groupings depend on the locations of the cities on each of the twelve components constructed in the factor analysis. The components in turn are a function of the variables selected for the study. "What factors come out of an analysis obviously depends on the composition and size of the set of variables actually used" (Janson, 1969, p. 316).

Table 25

King's Canadian City Groupings, 1951

1

Chicoutimi
Shawinigan
Trois Rivières

2

Sault Ste. Marie
Sudbury
Sydney
Timmins

3

Belleville
Brantford
Chatham
Cornwall
Guelph
Hamilton
London
Oshawa
Peterborough
Sarnia
St. Catherines
Welland
Windsor

4

Brandon
Lethbridge
Moose Jaw
Port Arthur
Regina
Saskatoon
Victoria

5

Calgary
Edmonton
Halifax
Kingston
Moncton
Montreal
Ottawa
Quebec
St. John's
Saint John
Toronto
Vancouver
Winnipeg

6

Drummondville
Granby
Sherbrooke
St. Jean
Valleyfield

7

Kitchener

SOURCE: Adapted from
King, 1966, Table II

Janson states that it is "deplorable" to use whatever variables are available because the number of variables from various fields determines the ranking and strength of the derived factors (Janson, 1969, p. 317). A significant problem of redundancy can also occur if the variables are not selected carefully to avoid this problem. It arises when one variable is contained in another or when two variables are parts of a total. This can easily happen when several percentages of one total are taken as the variables to be analyzed. Percentages of the same total tend to vary inversely with one another. They have high negative correlation coefficients, resulting in a component which accounts for this common variation.

King may appear to be guilty of selecting variables because of their availability. They are selected from census publications, from the Business Year Book, Canadian Municipal Financial Statistics, and the Atlas of Canada (King, 1966, p. 224). King does not explain the reasons for the selection of each variable. There is also considerable redundancy in his variables. Of the 52 variables selected, five are percentages or rates based on the age-sex structure of the population, eight are percentages of the labour force engaged in various industries, and eight others are percentages of the manufacturing sector engaged in specific types of manufacturing. These 21 redundant variables predetermine the most important components and consequently the groupings of the cities.

The answer to the question of the criteria for the grouping therefore, is that they are determined by the economic activities and age-sex structure of the cities' populations. The coincidence of a regional dimension in the groupings reflects the fact that industries and age-sex structures are not distributed randomly throughout cities but tend to be clustered in geographic regions.

Tracing the city groupings back through their components to the variables, we are able to more explicitly describe each group of cities. Group 1 contains industrial cities located in Quebec. The second group contains mining cities. The third group comprises small manufacturing cities in southern Ontario. They lack services which are provided by the large Metropolitan Areas near by. Group 4 includes smaller western Canadian cities with balanced economic bases. Group 5 consists of larger and/or older cities with diversified economic bases. Group 6 comprises smaller urban communities in Quebec located proximate to, and dominated by, Montreal and Quebec.

King's groupings may be useful in determining forces affecting the urban wage structure and its fluctuations. Focusing on the fluctuations, we note the number of cities in each group as compared with median residual change in average wage of male wage earners. The distribution is as follows:

	1	2	3	4	5	6
Residual gain	3	2	4	2	11	0
Residual loss	0	2	9	5	2	5

There are significant differences among the groups. Groups 1 and 6 both contain cities in Quebec but group 6 includes those smaller cities dominated by Montreal and Quebec. All five of these cities show negative residual change in average wage of male wage earners while the other three, Shawinigan, Chicoutimi, and Trois Rivières, show positive residual change in average wage of male wage earners.

Eleven of the thirteen large cities in group 5 show residual gain in average male wage. Only Saint John and Moncton show decline in this dimension. Both group 3 and group 4 cities predominantly decline in residual change in average wage of male wage earners. Both of these contain smaller cities in southern Ontario and the western provinces. King's groups reinforce earlier findings that larger cities tend to show positive fluctuations while smaller cities show negative fluctuations in average wage. They also provide the insight that within the province of Quebec, industrial cities show gain while smaller cities closer to Montreal and Quebec show declining residual change in average wage of male wage earners.

Turning to indicators of the established wage hierarchy, we note that the distribution of cities within groups according to their position relative to median migration impact on the labour force is as follows:

	1	2	3	4	5	6
Above median	0	1	8	4	8	1
Below median	3	3	5	3	5	4

This time cities in Quebec are not distinguishable with respect to the group they are in. Only Drummondville experiences above-median migration impact on the labour force. Mining cities tend to take positions below the median as well. Groups 3, 4, and 5 show approximately equal distributions with respect to migration impact on the labour force with about 60 percent of the cities falling above the median. King's groups fall into two large categories. Small Quebec centres and mining cities experience relatively less labour force growth from in-migration. Large centres and the smaller ones in southern Ontario and the West experience more labour force growth from in-migration.

A more valid indicator of the urban wage hierarchy than migration impact on the labour force is average wage of male wage earners in 1951. The distribution of cities in groups according to average wage of male wage earners in 1951 is as follows:

	1	2	3	4	5	6
Above median	0	3	11	1	7	0
Below median	3	1	2	6	6	5

All Quebec cities in groups 1 and 6 lie below the median. All of the smaller cities in the western provinces in group 4 also fall below the median. Port Arthur, which is in Ontario, is the solitary group 4 city lying above the median. Most of the small southern Ontario cities in group 3 lie above the median, confounding the association of city size and average wage. King's groupings reveal that smaller cities in Quebec and the western provinces lie below the median while smaller cities in southern Ontario lie above the median. The large cities in group 5 are distributed fairly evenly on each side of the median. The six below the median include Kingston and five others from Quebec and the Maritimes.

When the same variable, average wage of male wage earners, is taken at a later date, 1961, the distribution of cities in each group relative to the median is as follows:

	1	2	3	4	5	6
Above median	0	2	9	3	8	0
Below median	3	2	4	4	5	5

Quebec cities in groups 1 and 6 do not improve their positions. The most significant gain is in group 4 where two cities, Saskatoon and Regina, move above the median. One city, Montreal, moves above the median among the large cities in group 5. In group 3, Chatham and Brantford drop below the median. King's grouping indicates that Quebec cities in groups 1 and 6 continue to be disadvantaged. Smaller western cities show some gain while the smaller southern Ontario cities are redistributed somewhat lower in the 1961 wage hierarchy than they are in 1951.

King applies the same statistical procedures to the same set of cities using similar data for 1961. The same variables could not be obtained in all cases but the set of 54 variables that are used are easily comparable with the 52 used from 1951 data (King, 1966, p. 215).

He obtains different components in his factor analysis, suggesting that Canadian urban dimensions shifted over the ten year period. The first component is strongly identified with urban manufacturing, especially textile manufacturing, and it is identified with a weakly developed urban structure with respect to finance and service. The second component is identified with the Quebec population structure with predominantly young persons, fewer educated persons, and a smaller percentage of the population in the labour force. The third component is identified with metropolitan socio-economic structure - high population density, more immigrants, high median value of dwellings, number of through highways, and number of towns within a 50- and a 100-mile radius.

The fourth component appears to reflect the residential role of communities outside large Metropolitan Areas, particularly communities with high socio-economic status. The fifth component is identified principally with the service functions of many older and comparatively isolated communities. The sixth component indexes older, depressed cities. The pattern of loadings for the remaining components defies concise labelling (King, 1966, p. 219).

King again classifies the 106 cities into eleven groups. This time, the 46 cities studied here fall into nine of King's groups. The general pattern described by the 1951 groupings is preserved in the 1961 groupings, with few changes. The only significant changes occur to groups 3 and 5; the southern Ontario cities and the group of large metropolitan cities. Several southern Ontario cities move into other groups while a few form a new group. Group 5 loses all but Montreal and Toronto. The others join group 4, small western cities, or form a group of larger eastern Canadian cities. In spite of these changes the new set of groups retains distinct regional characteristics.

The 1961 groups appear in Table 26. Groups 1, 2, 6, and 9 contain so few cities from this study that they do not

Table 26

King's Canadian City Groupings, 1961

1

Chicoutimi
Timmins

2

Sault Ste. Marie
Sudbury
Sydney

3

Belleville
Brantford
Chatham
Guelph
Kitchener
Peterborough
Welland

4

Brandon
Calgary
Edmonton
Kingston
Lethbridge
London
Moncton
Moose Jaw
Ottawa
Regina
Saskatoon
Vancouver
Winnipeg

5

Cornwall
Drummondville
Granby
St. Jean
Trois Rivières
Valleyfield

6

Montreal
Toronto

7

Halifax
Hamilton
Quebec
Saint John
Sarnia
Shawinigan
St. John's

8

Oshawa
Port Arthur
St. Catherines
Victoria
Windsor

9

Sherbrooke

SOURCE: Adapted from King,
1966, Table IV

warrant discussion. Group 3 is smaller than it was in 1951 but retains similar properties and is geographically more concentrated. The cities involved typically have high scores on the fourth component, that of residential socio-economic status. Group 4 retains its western image but has attracted a number of cities from eastern Canada: Ottawa, London, Kingston, and Moncton. The member cities have relatively high rankings on the fifth component, the service function of older and comparatively isolated communities. Group 5, called group 6 in 1951, retains its strong regional Quebec character with the influx of Cornwall and Trois-Rivières. This group of cities ranks high on the first two components - manufacturing, shortage of services, and a typically Quebec population structure. Group 7 contains a group of relatively large eastern Canadian cities while group 8 cities are predominantly specialized manufacturing cities in Ontario. King admits to an inability to characterize them according to the variable loadings on components. For purposes of communication, group 3 shall be said to contain southern Ontario manufacturing cities, group 4, larger Ontario and western Canadian cities, group 5, smaller Quebec cities, group 7, eastern Canadian cities, and group 8, specialized manufacturing Ontario cities.

When cities in these groups are dichotomized according to the direction of residual change in average wage of male wage earners, the following distribution obtains:

	3	4	5	7	8
Positive change	3	8	1	6	0
Negative change	4	5	5	1	5

Group 8 is unique in that Victoria and all specialized manufacturing cities in Ontario experience negative residual change in average wage of male wage earners. A large proportion of small Quebec cities in group 5 experience negative residual change in average wage of male wage earners. Group 7, eastern Canadian cities, contains the largest proportion of cities showing positive residual change in average wage of male wage earners.

When the groups of cities are dichotomized according to their positions relative to median migration impact on the labour force, the following distribution obtains:

	3	4	5	7	8
Above median	5	11	2	1	1
Below median	2	2	4	6	4

Three groups contain cities that are predominantly below the median. These are small Quebec cities, eastern Canadian cities, and specialized manufacturing Ontario cities. The other two, western Canadian cities and Ontario cities, tend to experience substantial migration impact on the labour force.

A summary of the relationship between King's 1951 and 1961 groups and migration impact on the labour force can be stated thus: cities with above-median migration impact on the labour force tend to be large, and to be situated in western Canada or Ontario. Cities with below-median migration impact on the labour force include small Quebec cities, eastern Canadian cities, mining cities, and cities with specialized manufacturing functions.

When groups are dichotomized according to their positions relative to median average wage of male wage earners in 1951, the following distribution obtains:

	3	4	5	7	8
Above median	6	7	0	2	4
Below median	1	6	6	5	1

Groups 5 and 7 contain significantly large proportions below the median. These comprise small Quebec cities and eastern Canadian cities. Groups 3 and 8 contain significantly large proportions above the median. These comprise southern Ontario manufacturing cities and specialized manufacturing cities in Ontario.

In summary, King's 1951 and 1961 groups associate with average wage of male wage earners in 1951 in the following way: small cities within Ontario are above the median while small cities outside Ontario lie below the median; large cities in

Ontario and western Canada are above the median while those in Quebec and the Maritimes are below the median.

The distribution of cities within groups with respect to the median of the same variable, average wage of male wage earners, in 1961 is as follows:

	3	4	5	7	8
Above median	4	9	0	2	4
Below median	3	4	6	5	1

The distribution of small Quebec cities and eastern Canadian cities is preserved from 1951. They fall predominantly below the median. Similarly, the distribution of special manufacturing cities in Ontario is preserved. They fall above the median. The one exception in that group is Victoria. Group 3, southern Ontario manufacturing cities, experiences a shift below the median while larger western Canadian and Ontario cities shift to positions above the median compared to their 1951 distributions.

The association between both sets of King's groups and average wage of male wage earners in 1961 can be summarized as follows: groups that tend to concentrate above the median are those involving manufacturing cities in Ontario, larger, older cities, and western Canadian cities. Those falling below the median comprise cities located in Quebec and the Maritimes.

King's groupings reveal definite regional differences with respect to migration impact on average wage. If residual change in average wage of male wage earners can be construed as a measure of disequilibration of the urban wage hierarchy, King's data reveal that large cities, eastern Canadian cities, and small Quebec cities that are relatively independent of Montreal and Quebec experience positive disequilibrating change. Smaller cities across Canada and cities with specialized manufacturing functions show negative disequilibrating change.

Those cities positioned higher on the wage hierarchy tend to be located in Ontario and, if they are larger, western Canada. Cities in Quebec and the Maritimes, as well as smaller western Canadian cities, tend to be placed lower on the wage hierarchy. There is a tendency for smaller Ontario cities to shift downwards and for small western Canadian cities to shift upwards.

King's research shows that definite regional differences occur in the development of Canadian cities. His major variables reflect the association between industrial activity, age-sex structure and regionalism. The three are interrelated and deserve consideration in a comprehensive theory of differential urban development in Canada.

Capital Cities

Grossner suggests that capital cities in Canada tend to develop more quickly than others because of the expansion of government administration from 1951 to 1961. He found that seven out of ten cities showing an increase in economic specialization are capital cities (Grossner, 1970, p. 47). Table 24 appears to support his hypothesis. It contains nine capital cities, including Ottawa. Charlottetown and Fredericton are not large enough to be included in this study. Eight capital cities achieve residual gains in wages during the decade while one, Victoria, experiences a loss.

In Table 21, four cities lie above the median with respect to average wage of male wage earners in 1951 and Regina increases this total to five with respect to the 1961 data. Table 23 reflects the distribution found in Table 21. The same five cities lie above the median percentage change in labour force size attributed to migration as lie above the median for average wage of male wage earners in 1961.

While capital cities are evenly distributed in upper and lower segments of the urban hierarchy of wages, they experience more positive residual change in average wage of male wage earners than do other cities. Eight of the nine capital cities rank above the median in residual change in average wage of male wage earners as seen in Table 24. Only

Victoria experiences a small loss in residual change in average male wage.

Metropolitan Areas

It is interesting to note that the four western Canadian cities in Table 24 that experience negative residual change in average wage of male wage earners are without exception smaller than the six showing positive change. The same holds true for the cities in the Maritimes: all three cities showing residual decline in average wage are smaller than the two showing residual gain in average male wage. Although this clear relationship does not hold for Ontario and Quebec, it suggests a strong association between city size and residual change in average wage of male wage earners at the nominal level of measurement. The correlation coefficient of 1951 labour force size and residual change in average wage of male wage earners is 0.31, indicating a modest association at the interval level.

Metropolitan Areas represent one aspect of the dimension of size. They include all cities in Canada in 1951 with populations exceeding 100,000. Analysis of their positions on the urban wage hierarchy may provide some insight into the relationship between size and urban development.

Table 24 reveals that nine of the ten Metropolitan Areas achieve positive residual change in average wage of male wage earners. Only Windsor shows a loss. In Table 23 seven Metropolitan Areas hold positions above the median migration impact on the labour force while three, Winnipeg, Quebec, and Windsor, are below the median. In Table 21 eight cities are above the median average wage of male wage earners in 1951 with Montreal and Quebec falling below it. In 1961 in the same table, Montreal rose above the median leaving Quebec alone below it.

Metropolitan Areas offer more opportunities for agglomeration economies than do other cities in Canada. However they do not dominate the highest levels of the wage hierarchy. Apparently other forces also play an important role in this dimension of development.

Heartland-Hinterland

Maxwell (1965) attempts to classify by economic function the same set of 106 cities that King studies in his analysis. He obtains a functional profile of these cities using 1951 census data. The dominant functions include retail trade, transportation, manufacturing I, manufacturing II, government services, and extraction. Manufacturing was broken down into two groups because it included so many cities. Manufacturing I contains all cities where 50 percent or more of a city's excess employment is accounted for by manufacturing. All cities in manufacturing II are predominantly manufacturing cities, but this sector accounts for less than 50 percent of the cities' excess employment.

Maxwell's classification according to dominant function is displayed in Table 27. Manufacturing plays a much more significant role in cities located in southern Ontario than it does in any other region. High degrees of industrial linkages appear and a wide variety of industrial activities take place among these cities. A large population guarantees a convenient market for the manufacture of consumer goods. Its location makes it accessible to much larger markets in Europe and the United States. The existence within a relatively small area of energy, resources, raw materials, and port facilities contribute to rapid manufacturing development in that region.

The region with greatest initial advantages is the first to achieve sustained economic growth. Other advantages such as leadership in finance, education, research, and planning accrue to the heartland. Secondary manufacturing and service activity gravitate towards the heartland, leaving hinterland areas reliant on primary industries (Ray, 1971, p. 13).

Maxwell suggests that the Canadian urban system should be viewed as a heartland-hinterland dichotomy. Cities in the heartland enjoy excellent locational relationships for most manufacturing processes - the key urban function (Maxwell, 1965, p. 95).

Table 27

Cities Classified by Dominant Function

Retail Trade

Lethbridge

Transportation

Brandon
Moose Jaw
Moncton

Manufacturing II

Belleville
Calgary
Chatham
Chicoutimi
Edmonton
Kingston
London
Port Arthur
Quebec
Saint John
Saskatoon
St. John's
Winnipeg
Vancouver

Government Service

Halifax
Ottawa
Regina
Victoria

Manufacturing I

Brantford
Cornwall
Drummondville
Granby
Guelph
Hamilton
Kitchener
Montreal
Oshawa
Peterborough
Sarnia
Sault Ste. Marie
Shawinigan
Sherbrooke
St. Catherines
St. Jean
Sydney
Toronto
Trois Rivières
Valleyfield
Welland
Windsor

Extraction

Sudbury
Timmins

SOURCE: Adapted from Maxwell, 1965, Table V

The heartland is characterized by a great concentration of economic activity, financial power, and dense population. The region contains 52 percent of Canada's population in 1951 while comprising only 2.5 percent of the nation's land area. Cities in the periphery having to contend with the friction of space, must devote much of their economic energy to overcoming this friction.

Maxwell lays cities out on a grid which has one axis measuring excess employment in wholesale trade and the other measuring excess employment in manufacturing. He is able to identify several groups of cities according to their location on the grid. One group he describes as specialized manufacturing centres. These rank high on manufacturing and low on wholesale trade. Most of them are located in the heartland, which comprises the St. Lawrence lowlands and southern Ontario. These include the following:

- Trois Rivières
- Cornwall
- St. Jean
- Shawinigan
- Sault Ste. Marie
- St. Catherines
- Peterborough
- Kitchener
- Hamilton
- Windsor

- Granby
- Brantford
- Oshawa
- Drummondville
- Welland
- Valleyfield
- Guelph

A second group are called special cities. These are low on both axes. They comprise highly specialized government and extraction centres. This group includes the following four cities:

- Victoria
- Ottawa
- Sudbury
- Timmins

A third group are called regional capitals. These tend to perform central place functions. They rank high on wholesale trade and low on manufacturing. None of them are located in the heartland. They include:

- Saskatoon
- Regina
- Calgary
- Edmonton
- Saint John
- Brandon

- St. John's
- Lethbridge
- Halifax
- Moncton
- Moose Jaw

The final group are called metropolitan centres. They are relatively high on both the wholesale trade and manufacturing scales. The four cities in this group are:

- Winnipeg
- Vancouver
- Montreal
- Toronto

The remaining cities are multifunctional and do not fit into a specific category.

The heartland-hinterland framework may be useful in any theory of urban growth in Canada. Such a factor occurs when Ray (1971) factor analyzes 84 variables from the 1961 census for all cities in Canada with populations over 10,000. The variables with the highest factor loadings on this factor are percentage of the male labour force employed as craftsmen (+0.810) and percentage of the male labour force employed as machinists (+0.697). Such a framework touches on the basic assumptions of export base theory. Modifications of this theory should therefore be retained in spite of the poor showing of the location quotient coefficient in this study.

Economic Shadow

Another concept which should be integrated into a theory of urban growth in Canada is called the 'economic shadow'. It refers to the disadvantaged economic position of a city due to the strategic geographic location of others. Ray argues that regional differences in economic development in Canada are related to the regional impact of United States capital and entrepreneurial activity operating through an economic shadow effect (1968, p. 63). Regions that are advantageously located benefit from United States interests and cast a shadow over regions not so favorably placed. American branch plants are located in Canada near to parent companies in the United States from which they obtain complex parts, technical services, and supervisory advice.

Parent companies in the United States, looking to Canada for location of a branch plant, tend to base decisions of location on the usual laws of location such as distance from sources of raw materials and energy, proximity to markets and agglomeration advantages. At the same time the distance from the head office also is an important factor because a considerable amount of communication and transportation of personnel and goods occurs between the head office and the branch plant. Most United States manufacturing head offices are located in the eastern industrial belt, making the southern Ontario and the St. Lawrence lowlands the most advantageous location for branch plants. Thus the heartland casts an economic shadow over the remainder of Canada.

Within the Canadian system of cities, Toronto and Montreal cast economic shadows over all cities in Canada. They dominate communication and transportation of all parts of Canada with the eastern United States and Europe. Vancouver dominates economic interchange with Asia and with United States west coast industries. It casts an economic shadow over all cities in Canada dealing with corporations in Asia and the western United States. Within the prairie region, Edmonton and Calgary tend to dominate interaction with industries to the west while Winnipeg dominates interaction with industries to the east.

Financial Structure

The heart of a nation's commercial and business interests may be the complex of financial institutions including head offices of industrial, mining, retail, and other corporations, holding companies, chartered accounting, legal, and other business-service firms. These are concentrated in a few Metropolitan Areas. Most of the financial houses concerned with the working of the capital market, such as stock and bond exchanges, investment companies, and head offices of banks tend to locate in Toronto and Montreal (Putnam and Putnam, 1970, p. 79). Most insurance and trust companies tend to cluster in or near financial districts of large metropolitan cities.

A Financial Post survey in 1966 placed 38 percent of assets of corporations in Montreal and 37 percent in Toronto (Financial Post, 1966, pp. 25-26). Vancouver (6.5%), Calgary (5.0%), Hamilton (4.0%), Winnipeg (2.0%), and Quebec (1.0%) rank next in order.

Putnam and Putnam derive an "index of financial specialization" which reflects the financial dimension of economic activities carried out by twenty Canadian cities. The index is computed by dividing the percentage of the Canadian population residing in the city into the product of the percentage of total Canadian employment in financial institutions and percentage of the value of total cheques cashed in Canada.

Table 28 reflects the relative importance of various cities with respect to financial activity. It indicates the contribution that Toronto makes to the manufacturing cities located in Southern Ontario. It is noteworthy that the statistical manipulation by King showed these cities to be lacking in financial services. They apparently depend on Toronto for such services. The high rank of Regina, Winnipeg, and Calgary reflect their central place functions as financial centres for the prairies while Montreal plays this role for Quebec, Vancouver for British Columbia, and Halifax for the Maritimes.

Table 28

The Central Place Functions of Financial Institutions in Twenty Canadian Metropolitan Areas, 1961

Metropolitan Area	Index of Financial Specialization*
Toronto	7.93
Regina	4.42
Winnipeg	3.91
Calgary	3.82
Montreal	3.61
Vancouver	2.07
London	2.00
Edmonton	1.56
Quebec	1.43
Halifax	1.30
Victoria	1.24
Ottawa	1.19
Hamilton	1.18
Saint John	0.98
Windsor	0.97
Saskatoon	0.83
St. Catherines	0.63
Kitchener	0.54
St. John's	0.54
Sudbury	0.32

SOURCE: Putnam and Putnam, 1970, Table 3.15

*$S = \frac{ab}{c}$ where S = Index of Financial Specialization

a = Percentage of total Canadian employment in financial institutions.

b = Percentage of the value of all cheques cashed in Canada.

c = Percentage of Canadian population residing in that city.

Problems for Further Research

This dissertation has attempted to compare the applicability of three theories of regional development to urban development in Canada from 1951 to 1961. A number of research problems were brought into relief in the process of data collection and analysis which detracted from the work in this research project and which deserve consideration for further research.

An obvious problem which arises is the appropriateness of the selected indicator of each variable. Abstract variables are difficult to operationalize but when no traditional operational definitions have been established, the authenticity of one that is arbitrarily selected becomes more questionable.

A useful research project would be one in which possible indicators of entrepreneurial activity are evaluated. Such activity involves the establishment and/or expansion of firms, as well as the management of these firms with an aim of increasing profits. A major difficulty with operationalizing such a concept is the fact that profits are the result of numerous other factors while expansion of firms or establishment of new ones may reflect poor economic foresight rather than clever entrepreneurial activity. An operational definition of entrepreneurship would be valuable because the concept plays a crucial role in economic development theory.

The location quotient coefficient seems to be an inadequate measure of an economic base although tradition has entrenched it as the representation of the economic base in numerous research journals. Its weaknesses have been enumerated in Chapter 3 so they need not be repeated here.

The concept of urban development, as distinct from urban growth, requires elaboration. As a concept, urban development should extend beyond income of the urban area to include quality of life of the inhabitants of the urban area and its hinterland. Quality of life is a normative construct, like welfare and happiness, and so presents problems of clarity of definition and consensus (Wingo, 1973, p. 16). It is a useful concept, however, and deserves more attention, especially in the present context of concern for pollution, housing conditions, and traffic congestion. Social scientists should exert some energy towards the improvement of this concept and the refinement of an operational definition.

Given the bias in the conceptual definition of urban development, the operational definition also requires improvement. Average wage of male wage earners represents only a part of the income of the total labour force. It excludes female wages, a sector of the labour force which has recently grown rapidly. It also omits self employed persons. The most serious omission is the social dimension of development which necessarily accompanies sustained urban development. A mining town will remain a mining town, albeit a rich mining town, as long as no social, academic, health, or research institutions spring up to complement the economic advantages the town has accrued.

The concept of urban development should include a growth dimension. Growth is not a necessary condition for development but the two are usually associated. The measurement of urban growth is confounded by annexation if the annexed area contains a large population. Records seldom reflect the growth of the annexed region before it was annexed. Thus the analyst is required to construct an ad hoc model of growth for the annexed area or to treat the annexation as a sudden influx of migrants into the city in spite of the fact that the inhabitants of the annexed area may have moved there from the city itself or may have been born in the annexed area. This measurement problem could be nullified by improved record keeping. However, in the absence of such records, a model of growth of annexed areas would be useful for approximating the population of annexed areas for specific periods of time before the date of annexation.

Theories of regional development presume that economic development occurs before social development, if they make note of social development at all. This in itself is a questionable assumption, particularly under modern conditions in which selection of modern sites may be based upon the social development of competing regions. Modern forms of transportation and communication combined with government subsidies reduce the influence of the economic dimension on industrial location. Other considerations, such as availability of skilled labour becomes more important as manufacturing methods and technical skills become more sophisticated. The location of skilled labour is contingent on such factors as location of technical and academic institutions, climatic conditions, and availability of recreational facilities. It is not by chance that Canada's largest cities are located near large, and well developed recreational facilities. Toronto hugs the shore of Lake Ontario, not far from Georgian Bay; Montreal looks north to the Laurentians while Vancouver nestles between high mountains and the Pacific.

Amenities may play a larger role in determining the site of firms now than in the past but they are still constrained by economic impositions. Thus, the economic substructure continues to be paramount in any study of urban development. Regional economists have not yet established which industrial mixes are appropriate for sustained development of an area. Although a few industries such as petroleum and plastics, or steel and automobile manufacturing, are obviously complementary, recommendations regarding suitable industry mixes depends upon the resources, skills, and capital available within the region. Other factors to be considered include the proximity of markets, modernization of transportation and communication facilities, and the relative advantages of locating industries in one region _vis-à-vis_ another. The determination of growth-inducing industrial mixes for a specific region requires a careful analysis of advantages and disadvantages associated with that region. The industrial mix cannot be determined in abstract; it must be determined within the context of the attributes of each region.

Service industries play a major role in development at the urban, regional, and national levels. However, their contributions are difficult to evaluate. Under present conditions wherein services are beginning to dominate the economies of the more industrialized nations, regional economists need to know the contributions that various service industries make. A useful field of research would attempt to answer the question: what contributions do various services such as business services, legal and financial services, and personal and social services make to the development of a region or city?

Urban development is a dynamic concept. It occurs as a result of the flow of goods and services out of the area as well as the result of the flow of capital and skills into the area. Another essential component of development is the flow of ideas between urban areas and their dispersal into the urban hinterland. The discipline of regional science would benefit from the construction of models of these flows.

The urban hinterland plays a crucial role in growth pole theory. Spill-over effects from the growth centre supposedly generate growth in the surrounding region. This assumption influences policy decisions of governments in their selection of towns as sites of new industries. These towns are supposed to become growth centres which generate development in the surrounding depressed regions. Apparently Toronto generates more spill-over effects than Montreal (Higgins, 1972, p. 177). This condition has adverse effects on the economic development of the province of Quebec.

A thorough analysis of conditions that constitute spill-over effects is needed. Under what regional and economic conditions do they occur? Which types of spill-over effects are more beneficial than others? This requires an analysis of the types of linkages that exist among large cities and smaller towns nearby.

A final suggestion for future research brings into focus the part played by the region in the development of an urban area. Regional economists have not established what industrial mix is necessary to make a town a growth centre. The appropriate industrial mix probably depends upon the economic and social attributes of the surrounding region. This implies that the definition of a growth pole depends upon the region in which it is located. A growth pole in the prairies differs from one in the Maritimes, Quebec, or southern Ontario. Research into the interplay between the centre and the region would contribute to Canadian economic and social development as well as to growth pole theory as a theory.

Summary of Chapter 5

The first section deals with the major variables associated with each of the four indicators of urban development: average wage of male wage earners in 1951, average wage of male wage earners in 1961, residual change in average wage of male wage earners, and migration impact on the labour force. Migration impact and average wage for the two census years are intercorrelated but residual change in average wage does not correlate with the other three. This suggests a model of a stable urban wage hierarchy, as indicated by the three intercorrelated variables, with some destabilizing influences, as indicated by residual change in average male wage.

Factors which contribute to stability of the urban wage hierarchy are profits, agglomeration, and fluctuations in the export base. Factors which tend to disrupt the hierarchy include fluctuations in profits, proportion of the labour force in manufacturing and service industries, agglomeration, and profits in 1951. The causal relationship between residual change in average male wage and each of the independent factors is difficult to explain. It is surmised that the associations may reflect external factors such as inflation during the early 1950's and shifts in economic activities which might contribute to disequilibrium of the wage hierarchy.

The variables included in this study are selected from three theories of regional development: export base theory, sector theory, and growth pole theory. They do not encompass all forces which contribute to Canadian urban development. The second section of the chapter surveys some other factors which have been considered in the literature.

Capital cities and Metropolitan Areas rank high on the urban wage hierarchy. Capital cities benefit from expansion of provincial and federal government administration. Metropolitan Areas expand because of advantages of agglomeration and social overhead which they have over smaller cities.

An important factor in differential urban development in Canada is regionalism. Ontario and British Columbia rank highest with the prairies, Quebec and the Maritimes following in that order. The cities investigated here appear to reinforce this regional disparity. The nature of the cities' economic activities, which tend to be homogeneous within regions, may help to explain these regional differences.

Southern Ontario cities rank high on the urban wage hierarchy. One reason for this can be expressed in terms of a heartland-hinterland framework. Southern Ontario had initial advantages of proximity to large markets, ports, and access to energy and raw material sources. A concentration of economic activity, financial power, and dense population evolved, leaving hinterland areas largely dependent on primary industries. Cities in the less dense hinterland were forced to devote much of their economic energy to overcoming the friction of space.

Cities in southern Ontario have another advantage in that they lie nearest the industrial belt in the eastern United States. Parent companies prefer to set up branch plants in southern Ontario where advantages of proximity to the head office complements the other locational advantages of that region. Cities in the heartland, particularly Toronto and Montreal, consequently cast economic shadows over less advantageously located cities in Canada.

Chapter 6

Demographic Dimensions of Urban Growth

The intent of this chapter is to provide some insight into the demographic and ecological changes that a North American city undergoes while it is growing. Urban growth may differ from urban development, i.e., a city may grow (increase in population) with development or may develop (increase in per capita income) without growing; however such conditions are very unusual. In almost all cases, growth accompanies development. As incomes increase, migrants are attracted to the area and potential out-migrants choose to remain. As population increases, demand for goods and services attracts industry which provides jobs and income for the resident labour force.

The causal arrow connecting demographic growth and economic development in the general context must therefore have two heads. Growth contributes to development and development in turn results in further growth. Growth and development do not continue ad infinitum of course. Finite resources and capital, and diminishing returns act as checks to infinite growth. The competition of other cities limits the available capital and resources likely to be attracted to a single city. The friction of space dampens the attraction of people and industries from great distances. While growth and development complement each other, they also counteract one-another. Additions to the labour force pull down wages. Production increases cause supply to exceed demand, reducing the prices of products and lessening profit. Declines in wages, prices, and profits result in lower per capita income, making the area less attractive to future migrants.

The interaction of demography and economics is complex. The following sections deal with some aspects of economic change under conditions of population growth, patterns of urban growth and density, demographic differentials of small and large cities, growing and stagnant cities, and an evaluation of the three theories of regional development in terms of demographic realities.

Economic Change Under Conditions of Population Growth

Demographic change affects the economy through its effect on productive capacity and on changing patterns of expenditure. Productive capacity is a function of labour force size and structure. Labour force size depends on labour force participation rates. Birth rates have a delayed influence on labour force size but birth rates are themselves influenced by contemporary female labour force participation rates. A low birth rate in an area frees women to participate more fully in the labour force.

Migration has a more direct impact on labour force size. It also affects the structure of the labour force in that migrants are predominantly younger than average workers in the area. Migration may occur to meet the demand for special skills not already available in the area. It may change the structure of the labour force in terms of skills, sex-ratio, and education, as well as in terms of age distribution.

Changes in labour force structure and size, whether they be the result of migration, participation rates or birth rates of an earlier decade, not only affect the productive capacity of the economy but also alter patterns of expenditure. An increase in the labour force because of earlier birth rates or contemporary in-migration results in a larger population which increases demand for educational services, housing and public utilities. A larger labour force will cause a shift in the overall consumption function, which defines the consumption expenditures that take place at all relevant levels of national income, while changes in the structure of the labour force affect the prevailing pattern of expenditure (Denton and Spencer, 1970, p. 4).

A more thorough analysis of the relationship between population growth and productive capacity suggests that under some conditions population growth can have adverse affects on productive capacity. A popular view since Malthus has been that continued population growth results in a drain in resources, reduced capital accumulation and a greater strain on the organizing capacity of the society (Kuznets, 1960, p. 325). However Kuznets proposes three reasons why productivity per worker would be greater in an increasing labour force assuming adequate availability of capital (Kuznets, 1960, pp. 326-330).

1. Assuming the existence of unexploited resources, additions to the labour force permit greater and more efficient utilization of the resources. If this increased utilization is combined with a more specialized division of labour, the result is almost certain to be an increase in per capita production.
2. A growing labour force will have relatively larger groups of young workers than will a stagnant labour force. Younger workers are more mobile and therefore more likely to move into new regions and growing sectors of the economy. If labour is not mobile, new and expanding industries are faced with the problem of increasing wages in order to attract workers. This raises the cost of the product and reduces profits, thus lowering the ceiling on potential growth.

Immobility of the labour force is also disadvantageous in cases where old industries are deteriorating. When the labour force continues to stick to declining industries, the probability that per capita product will rise is decreased. A growing population will have proportionately more young workers in the labour force. Their mobility can be a factor in increasing the productivity of the economy.

3. A large population guarantees a larger absolute number of competent talented persons. Productivity depends to a large extent on technology. Technology requires techniques and technologists as well as machines. Ideas are proposed, evaluated, and applied. Thus imaginative persons must develop inventions, entrepreneurs must evaluate and apply them while skilled tradesmen and administrators must work with the machinery and organization to ensure the success of the innovation. Ideas need not stem from internal sources; they may come from other countries, but the technological expertise is still necessary if they are to be successfully inaugerated. If the ideas are developed by local inventors, the local industries are likely to benefit from them before their competitors, giving them an advantage which they would not have if they were forced to adopt the innovations of their competitors.

Kuznets appears to confuse a large population with a growing population in this third point above. If a population is large, then it need not grow in order to have a large absolute number of competent, talented persons.

The assumption Kuznets makes of an adequate supply of capital during population increase should be critically analyzed because it contradicts the popular notion that population increase hinders capital accumulation. This is the reason proposed by many experts why underdeveloped countries are experiencing difficulty in their economic progress. It plays a crucial role in a number of models of economic development proposed for underdeveloped countries. A sacrifice in consumption is necessary in order to increase capital formation. However, if consumption is increasing, as it ordinarily is when the population size is increasing, the task of converting some consumption to investment becomes more difficult.

Capital formation tends to accompany population growth under only a few favourable circumstances, these being the existence of unexploited resources and an area which is not heavily populated (Friedman, 1960, p. 347). The presence of unexploited resources reduces the rate at which diminishing returns occur. The shortage of population means that increases in that population allow greater division of labour, expansion of knowledge, and an increase in the size of the domestic market, allowing economies of scale to increase productivity and encourage further investment in spite of rising consumption.

Moving from the supply to the demand side of economics, we find that an increasing population may have positive or negative effects on the economy because of the increase in demand for goods. Demographic variables affect aggregate demand (and consequently national income and employment or prices) by affecting (Coale, 1960, p. 352):

1. the consumption function;
2. net private investment;
3. government expenditures on goods or services.

The consumption function, as stated earlier, defines the consumption expenditures that would take place at all relevant levels of national income. The relationship between change in population size and the consumption function needs little elaboration. If a family has more children, it will obviously need to consume more food and clothing, and will tend to find a larger home and car if the original ones are not large enough.

/ Private investment will tend to increase under the condition of a growing population because of the increased demand for goods. Such investment may increase on the basis of an anticipated increase in demand. Even if total national income does not increase, investment will be made in those sectors of the economy that are particularly sensitive to numbers of people rather than to their life styles or living standards. These include food consumption, household necessities, and children's clothing. In addition, investors tend to be reasured in most other sectors of the economy under conditions of population increase.

Government expenditures will also tend to change in accordance with population change. Expenditures will increase to provide the same standards of government services to more people. Expenditures will also be made in an attempt to meet anticipated population needs.

In summary, population growth will tend to have favourable economic effects only under conditions of excess resources and under-population. Changes in technology or in social organization may compensate for depleted resources or over-population but ignoring these possibilities, the preceding statement is a fairly adequate generalization. Under these two favourable conditions, an increasing population will increase productivity because the labour force will be more responsive to jobs in different regions and in different sectors of the economy. It will encourage investment because of anticipated higher demand for commodities and it will increase government expenditure to provide services to the larger population.

Migration

Urban growth, which has become one of the major social phenomena of our time, could not have come about at the rate it has by natural growth alone. It has required the movement of people from rural to urban areas and from smaller urban areas to larger ones. Differential migration is a crucial factor in differential urban development, and at the same time, urban development promotes in-migration. Migrants can be expected to move not only to areas with high wage levels but also to those areas experiencing an increase in wage levels (Greenwood, 1973, p. 96).

According to the simplified law of supply and demand, the movement of labour from one urban area to another should lower wages in the recipient area and raise them in the sending region. However interaction between level of income and migration is much more complex than this. Migrants come equipped with specific skills and it is the wage levels of occupations requiring those specific skills that are subject to depressing forces. Income levels in related industries may rise if these industries benefit from the lowered wages of the industries receiving the migrants. They benefit because they are able to purchase products from the industry hiring migrants at prices made lower because of the decrease in cost of labour. The cost of labour may be reduced not only because of lowered wages but also because added workers allow economies of scale and finer division of labour.

Another set of industries that benefit from the influx of migrants although they are not linked to the recipient industry are ones that provide services and commodities to the migrants. New residents of an area require food, housing, furniture, clothing, transportation, and recreation just as permanent residents do. The increase in demand for these commodities adds to the income of the relevant industries. This increase may attract investment which creates jobs, attracts more migrants, and adds to the total income of the area.

The preceding paragraphs imply that recipient areas of migrants achieve higher levels of economic development because of economies of scale and the in-flow of capital that accompanies migration. In-migration of labour force members not only causes greater employment growth, but also induces income growth in urban areas (Greenwood, 1973, p. 108). However, Perloff *et al*. warn that

> large numbers and high rates of population growth have quite different effects under different economic circumstances. Where job opportunities are expanding rapidly, population growth can itself be a force for growth in *per capita* income; but where job opportunities are not expanding at all or expanding only slowly, it may well have an opposite effect (1960, p. 589).

Louis Parai concurs with this view (1974). Increase in per capita worker, which is the essential indicator of economic development, requires concomitant increases in the other factors of production. Such increases result in better allocation of resources, economies of scale, or changes in technology, leading to increasing efficiency of production. Only when existing stocks of capital are being underutilized because of a labour shortage can in-migration alone add to the output of an area (Parai, 1974, p. 24).

One of the major factors in differential migration is economic. Migrants move because they wish to improve their economic condition. This suggests that any predictions of migration are conditional upon economic trends. However, some models have been proposed which ignore economic cycles. One such model is that proposed by Zipf (1949) which relates quantity of total migration between two centres to the sizes of the centres and inversely to the square of the distance between them. The total migration is thus:

$$M = \frac{P_1 \; P_2}{D^2}$$

where M = total migration,

P_1 = population of one centre,

P_2 = population of the other centre,

D = distance between the two centres.

Zipf's model is analagous to Newton's Law of Gravity. It suggests that interaction between two places is a random event depending on the number of potential actors and the distance between the centres. Isard suggests several criticisms of the gravity model when it is applied to inter-metropolitan migration (1960, pp. 504-512). It ignores variation in the shape and geographical size of the two interacting centres. Variations in population density and distributions of this density are also ignored. Such omissions may be serious when the model is applied to regions, but not so important when applied to large urban areas.

A more relevant criticism is the use of total population as a measure of mass. When migration is being estimated, a more relevant measure of mass would be labour force size or total income of centres. Similarly measures of distance that are better than straight-line mileage would be road mileage or cost of travel between the two areas. An adequate operational definition of social distance would be most appropriate - one that takes into account the similarity of cultural patterns, degree of knowledge that inhabitants of one centre have of the other, and the extent to which social and economic roles coincide in the two centres.

Another method of improving the gravity model is to provide weights for various segments of the population. Not all members of each population have the same propensity to move. If a weighting procedure could be devised which would account for these differences, the model would adhere more closely to empirical evidence. One suggested weighting procedure is to multiply the population of each centre by its average per capita income. Other possible weights might be indicators of occupational structure or sex and age composition.

Finally, the question of what exponent should be used with each variable needs to be resolved. The exponent of each population is typically unity while that of the distance between the centres is two. Much more evidence needs to be collected to determine what conditions dictate these particular exponents and which exponents are appropriate under other specified conditions.

A second model proposed by Stouffer recognizes the depressing effect of occupational opportunities that exist between the two centres. His model of intervening opportunities is that

> the number of people going [a distance S] from a point is directly proportional to the number of opportunities on the perimeter of a circle with radius S and inversely proportional to the number of opportunities on or within that circle (Stouffer, 1960, p. 1).

Stouffer's model is cognizant of economic forces in that economic advantages to migrants are manifested in terms of job opportunities. Zipf's model is less easily integrated into an economic perspective on migration, but it can be adapted to this view because larger urban areas offer more opportunities to potential migrants than smaller urban areas.

Both models complement an empirical generalization that in the United States there is a net migration flow upward in the urban hierarchy. Persons born in small places tend to move to larger places replacing migrants who have moved to even larger places (Hudson, 1972, p. 1).

Stouffer's model differs from Zipf's in two significant ways. He defines distance in terms of the cumulated sizes of intervening centres. This represents an approach to measuring social space in that distance is a function of the number and size of places an individual must pass through in travelling from one centre to another. He also specifies an origin and a destination of migration flow, thus specifying the direction of the flow.

Lowry developed a gravity model that attempted to use variables other than migration to help "explain" migration (1966). His model took the form:

$$M_{i \to j} = K \left[\frac{U_i}{U_j} \cdot \frac{W_j}{W_i} - \frac{L_i L_j}{D_{ij}} \right]$$

where:

$M_{i \rightarrow j}$ = number of migrants from centre i to centre j;

L = number of persons in the nonagricultural labour force;

U = unemployment as a percentage of the nonagricultural labour force;

W = hourly manufacturing wage;

D_{ij} = airline distance from i to j in miles.

The causal interpretation is that people migrate from lower to higher wage areas, and from areas of surplus labour to areas with labour shortages, and that attraction between areas is a function of their sizes and the distance between them.

Lowry applied his model to 1960 United States Census data on migration in 90 SMSA's. After elaborating the model to distinguish between civilian and military members of the labour force, he found that multiple regression of the independent variables on migration flow resulted in all but two of the regression coefficients being significantly different from zero at the 0.01 level of confidence. The two variables that failed to explain significant variation in migration flow were wage level and unemployment level of the sending SMSA. It appears that volume of out-migration is independent of the prevailing labour-market conditions of the SMSA.

An explanation of the inability of these conditions to predict out-migration is suggested by Perloff, et al. (1960, p. 590). A region with high wages and numerous job opportunities attracts migrants from all parts of the country. But out-migration from an over-populated area is a function of personal decisions within that area. These personal decisions are based on a number of other factors besides wage levels and unemployment rates. Thus the labour-market conditions are not statistically associated with out-migration from a sending area but they are associated with in-migration into a receiving area.

Lowry found that out-migration tended to be a function of city size rather than economic conditions. He compared San Jose with a history of accelerating growth in employment, and Albany, with a declining rate of growth (Lowry, 1966, p. 23). San Jose experienced net in-migration of 129,000 while Albany experienced net in-migration of 19,000 persons. Yet, gross out-migration for the two areas was approximately the same: 74,000 for San Jose and 71,000 for Albany. These values reflect the similarity in size of the two areas; the 1960 populations were 642,000 in San Jose and 658,000 in Albany. Regression analysis of out-migration on various independent variables revealed that labour force size is the only one that seems to affect out-migration (Lowry, 1966, p. 25). A second variable is age-sex structure because migration is age - and sex-selective but this variable gains prominance only when there is considerable variation in the age-sex structure of the urban community.

The net effect of migration for a prospering area is an increase in population. In-migrants tend to be young and vigorous and although some move on, many remain. The depressed area loses its younger working population. Older persons tend to remain as do the less vigorous of the younger population. The remaining labour force declines in quality and becomes less mobile on account of its age-structure and the selective effects of migration.

Lowry set up a second model (1966, p. 38):

$$dM_i = a_0 + a_1 dP_i + a_2 dQ_i + a_3 dA_i + a_4 dE_i + a_5 dI_i$$

where:

dM_i = net change in population 15 to 64 years of age attributable to migration;

dP_i = net change in the number of residents 15 to 64 years of age attributable to natural increase;

dQ_i = net change in civilian nonagricultural employment;

dA_i = net change in military personnel;

dE_i = net change in school enrolees 14 to 29 years of age;

dI_i = change in median income for families and unrelated individuals.

Lowry states that this model is useful for predicting net migration in place i because it can be done without direct comparison of the characteristics of place i to those of other places. "At most we need to consider the overall mobility of the population of the United States (a function of its size and structure) and the condition of the national labor market" (Lowry, 1966, p. 35). The logic of his explanation escapes this student of migration models. The variables he uses in his regression model reflect population change and wage change. Relationships between net migration and population change are at least partly tautological. Net change in employment is almost certain to vary with net migration, yet the use of the regression model implies that net change in employment has a causal affect on net migration. Lowry recognizes this problem, but resolves it by defining change in the labour force, which is the change from 1950 to 1960 in absolute numbers of the number of persons employed, as "change in employment opportunities" (1966, p. 42). He is thus able to set up a causal relationship between change in employment opportunities and net migration.

All but one variable had coefficients significantly different from zero at the 0.05 level of confidence. The coefficient of net change in school enrolees 14 - 29 years of age did not quite meet this test.

The independent variables account for virtually all of the variance in the dependent variable ($R^2 = 0.99$), as would be expected because one variable alone, change in the labour force, has a correlation coefficient of 0.98 with net migration. It is worthy of note that the regression model is subject to multicollinearity. Net change in civilian nonagricultural employment and net change in school enrolees 14 - 29 years of age have a correlation coefficient of 0.92.

Urban Size

Most empirical studies do not isolate urban size as a crucial factor in urban growth. One reason for this is that growth is usually measured as a percentage of a base population at an initial point in time, which places large cities at a disadvantage because they must amass a much greater absolute growth in order to compare favourably with the percentage increases of smaller centres.

Large urban areas are also at a disadvantage if development is measured in terms of income change because they are not likely to experience the extreme income changes of smaller centres that happen to be dominated by a single industry with rapidly expanding markets. Large cities are likely to include new, fast-growing industries, older, slower-growing industries, and some decaying industries that act to dampen income increase. This industrial mix assures the large city of long-term continued growth but prohibits it from achieving rapid economic development in any given decade (Thompson, 1968, p. 52).

The larger urban areas in any system of cities have a number of advantages over their smaller counterparts. Most of the inventions and innovations would tend to come from their ranks not just because of the greater absolute numbers of potential inventors and entrepreneurs, but also because of the mix of ideas in larger centres. If one man in a city of 10,000 produces a good idea per year, how many ideas are produced in a city of a million through the interaction of the ideas of one hundred men? The larger centre is guaranteed a fairly regular output of innovations, while the small centre must wait for several years for a good idea that enhances its economy.

The large centre tends to produce proportionately more innovations not only because of the greater potential for the exchange of ideas but also because the innovators and entrepreneurs are attracted to the major city and away from the smaller one. The metropolis houses the large university, the research centre, and the technical school, providing funds, facilities, and personnel for the development and testing of new ideas.

The innovations are often applied in the large city first because they originate there and because larger corporations, which can afford to experiment more readily than the small business, are also located in the metropolis. The large urban area has the largest and most varied supply of capital with which to operationalize the invention (Thompson, 1968, p. 54). The other essential component, labour, is also located in the larger area, which contains a greater diversity of skills as well as a larger absolute labour pool.

The growth of large cities is more stable and consistent than that of smaller cities because they contain a greater diversity of industries. The fast growing industries will tend to balance the slower growing, stagnant or declining industries, resulting in a stable pattern of growth over extended periods of time. Diverse industrial mix represents one of several advanatages that large centres have over smaller ones in their competition for economic dominance. Others include a concentration of ideas which spark inventions and innovations as well as reduced friction of space for the combustion of these ideas. A large labour pool guarantees a regular and skilled supply of labour. A large market close at hand and the presence of potential linked industries provides agglomeration economies for existing firms and encourages other firms to locate in that area. The existence of diverse, inexpensive services and facilities, such as public transportation, power sources, financial institutions, and research facilities, also make larger cities much more attractive than smaller ones to firms looking for places in which to expand.

The association between industrial mix and urban size in the Canadian context has been investigated independently by Crowley (1973) and by Marshall (1975). Crowley calculated coefficients of correlation for city size and six indices of specialisation/diversification. He defines specialisation as "the extent to which the labour force is concentrated in a small number of industry groupings" and diversification as the extent to which the labour force is dispersed over industry groupings, making the concepts opposing points on a continuum (1973, p. 92).

The first measure of specialisation (S_1) used by Crowley is a simple one. If all industries in each city are ranked from largest to smallest, the measure can be computed as follows:

$$S_1 = \frac{\sum_{i=1}^{3} x_{ij}}{\sum_{i=1}^{n} x_{ij}}$$

where: x_{ij} = employment in industry i in city j;

n = the number of industries.

It indicates the proportion of a city's total labour force occupied in its three largest industries.

The second measure of specialisation is equally simple. It is designated S_2 and is the reciprocal of the number of industries required to account for 80 percent of the labour force.

The third measure of specialisation, S_3, is the Gini index which results from plotting the successive accumulation of X_{ij} from 1 to n. It is an improvement on the first two measures in that it is not computed on the basis of an arbitrary demarcation value, e.g., three largest industries, or 80 percent of the labour force. The Gini coefficient yields the proportion of the total distribution which varies from a hypothetical one with equally populated industries.

S_4 is the sum of the squares of all location quotients greater than one. The location quotient is defined as:

$$\frac{e_i/e_t}{E_i/E_t}$$

where e_i = employment in industry i in a particular city;

e_t = total employment in a particular city;

E_i = employment in industry i for all cities combined;

E_t = employment in all industries for all cities.

If any industry in a city absorbs an above-average proportion of the city's labour force, the location quotient is squared and all such squares are summed.

The fifth measure of specialisation, S_5, is calculated as:

$$S_5 = \Sigma\left({}^{e_i}/_{e_t}\right)^2$$ for all values of the location quotient greater than 1.

This measure can be interpreted as the sum of the squares of the proportions of all industries containing an above-average percentage of the city's labour force.

The sixth measure of specialisation, S_6, is the same measure used in this thesis (see pages 85-92). Crowley computed these six different indices of specialisation/diversification because he was concerned with the lack of association of urban size and industry specialisation. He believed that weak associations in previous studies stemmed from the choice of indicators of specialisation (Crowley, 1973, p. 91). He illustrates his point by setting up a matrix of correlation coefficients of the six indicators using 1961 Canadian census data.

The matrix is presented in Table 29. Indicators 1, 2, 3, and 5 are highly intercorrelated, with coefficients of correlation of 0.87 or better. This suggests that the four indicators are interchangeable, but one must remember that there is still from 10 to 20 percent variation in each measure that cannot be accounted for by the others. Indicators 1, 2, and 3, differ from the others in that they are computed using data that pertains to the city alone. No information about other cities is used. Indicator 5 uses some information about other cities, but not as much as 4 and 6, which measure the degree of specialisation of a city relative to all others. Indicators 4 and 6 have a relatively high correlation coefficient, 0.75, but not high enough to make them strictly interchangeable.

Crowley then calculated the correlation coefficients of the six indicators and population and labour force size of Canadian cities in 1951 and 1961. The coefficients are presented in Table 30. The most obvious lessen from the Table is that regardless of the indicator used, the degree of specialisation/diversification is not closely associated with urban size. The second point to note is that indicators 1, 2, 3, and 5 are more closely associated with urban size than are indicators 4 and 6. The last two indicators attempt to measure specialisation/diversification relative to other cities while the other four measure specialisation/diversification as an internal phenomenon, without recognizing the distribution of the labour force in other cities.

The third point to note is that the strength of the association between specialisation/diversification and city

Table 29

Coefficients of Correlation Among Various
Specialisation/Diversification Indices*

	S_1	S_2	S_3	S_4	S_5	S_6
S_1	-					
S_2	0.87	-				
S_3	0.91	0.95	-			
S_4	0.58	0.41	0.42	-		
S_5	0.98	0.90	0.95	0.51	-	
S_6	0.33	0.16	0.22	0.75	0.36	-

*See text for description of indices.

SOURCE: Adapted from Crowley, 1973, Table 1.

Table 30

Coefficients of Correlation for City Size and Specialisation/Diversification Indices*

	Population		Labour Force	
	1951	1961	1951	1961
S_1	-0.52	-0.36	-0.59	-0.49
S_2	-0.64	-0.52	-0.70	-0.59
S_3	-0.57	-0.39	-0.63	-0.49
S_4	-0.24	-0.20	-0.31	-0.31
S_5	-0.53	-0.38	-0.60	-0.49
S_6	-0.01	-0.11	-0.06	-0.05

*See text for description of indices.

SOURCE: Adapted from Crowley, 1973, Table 2.

size declined from 1951 to 1961 and the fourth point is that coefficients are higher for the labour force size than for population. The size of the labour force is more an economic dimension of size than is total population which is also a reflection of the age distribution. Labour force size more directly reflects the capacity of production of an urban area.

The question that bothers Crowley, and which he is unable to answer is why the association of specialisation/diversification and urban size is not stronger. He suggests that one answer is the inadequacy of the measures of specialisation/diversification, as indicated by the variety of correlation coefficients in Table 29. He does not investigate any other possibilities.

Marshall establishes the weak association of urban size and specialisation using different indices and then explores the possible intervening variables that may be the cause of attenuation (1975). The two indicators he used were the Gini index and the mean absolute deviation index. The mean absolute deviation index, M_j, for N industries and j cities is:

$$M_j = \frac{\sum_{i=1}^{N} |x_{ij} - \overline{x}_i|}{N}$$

The index is simply the mean value of the absolute differences between the industry percentages in a particular city and the corresponding percentages in the national average. Using rank correlation, he found Spearman's rho to be 0.421 for city size and the Gini index, and 0.452 for city size and the mean absolute deviation index (Marshall, 1975, p. 41). He investigated those cities which deviated furthest from the regression line and consequently contributed most to their expected rank and found that those with positive deviations (they were more specialised than predicted on the basis of size) were characterised by being dominated by one, or at the most two industries. A disproportionate number of them were located in southern Quebec where there was a smaller variety of industries and a greater dependence on labour-intensive rather than on capital-intensive industries (Marshall, 1975, p. 47). Those with large negative deviations (more diversified than predicted on the basis of size) performed regional service functions. They were central places and tended to be located in the prairie provinces.

Another possible factor that reduces the inverse relationship between degree of specialisation and urban size is the fact that large urban agglomerations do specialise in some industries. Financial institutions and head offices of most corporations and business-service firms are located in Metropolitan Areas. These are what Stone calls "metropolitan functions" (1967, p. 187). Duncan et al. suggest that the fields of finance, wholesale trade, and business services are attracted to large urban centres (1960, p. 11). These sectors benefit from two advantages larger cities have over smaller ones: size and potential interaction. Wholesale trade can be expected to depend heavily on economies of scale provided by the large populations while finance and business services would utilize opportunities for large flows of information. Table 28 reflects the strong association between the degree of financial specialisation and urban size. Hartwick and Crowley investigated the location quotients of several industries they believed to be associated with urban size (1972, pp. 237-245). They found "finance, real estate, and insurance", and "trade" to be positively associated and "services" negatively associated with urban size.

Stone computed location quotients for industry groups he considered to be sensitive to the performance of metropolitan functions (1967, pp. 189-196). He selected (a) wholesale trade, (b) finance, insurance, and real estate, (c) business services, and (d) fabricating industries as those most likely to be found concentrated in large urban centres. He then ranked cities according to their location quotients in these four industries.

Toronto and Montreal ranked highest in the hierarchy of Metropolitan Areas with regard to concentration of economic activity in the performance of metropolitan functions. Behind them was a cluster that comprised Vancouver, Calgary, Winnipeg, and London. In the third level of the hierarchy were Edmonton, Saint John, and Windsor, while Sudbury and Victoria ranked lowest of the remaining eight Metropolitan Areas. This hierarchy reflects a fairly persistent association between urban size and degree of specialisation in metropolitan functions.

With the exception of these particular industries the association of industrial diversity and urban size would probably be much stronger. The broad labour force base encourages the location of a greater variety of economic activities that benefit from the agglomeration economies offered in a large urban centre. A diversified industrial structure precludes the likelihood that one industry will dominate the labour force, whereas in smaller centres a single growth industry can make a much larger impact. At the same time, the growth industry in the small town will continue to dominate because service industries are not so readily attracted to the small town. They more willingly move to a large centre where they are guaranteed a larger market and a more dependable one should that one growth industry begin to decline. They can be fairly sure that other growth industries will replace the original should it fail.

Urban economists have given little attention to the relationship between urban size and average per capita income. One would expect some association since the larger cities achieved their size through the attraction of migrants as well as through natural increase. In an economy where workers are free to move, they will tend to select those areas which promise higher real incomes. Larger cities are supposedly more promising places for industries to locate because of agglomeration economies, and one would expect the firm to pass on some of these profits to the workers, if for no other reason than to attract and retain better workers.

Table 21 indicates that there does seem to be some association between size and average wage of male wage earners in Canadian cities. All of the larger Canadian cities with the exception of Quebec and Montreal rank above the median in both 1951 and 1961. In his study, Edwin Mansfield makes the probabilistic conclusion that the chances are "about 1 in 3 that the 1949 median income in a randomly chosen city under 5,000 will be higher than in a standard metropolitan area of 100,000 to 250,000" (1957, p. 317). The difficulty in drawing taut conclusions regarding the association of income with urban size is that there are so many exceptions. Although the average income for small cities is less than that for large cities, the cities with the highest per capita income are often very small. Witness the cities in Table 21 with the highest average wage. In 1951 none of the six highest ranking cities were Metropolitan Areas. In 1961 the Metropolitan Areas make considerable gains but seven of the top twelve have populations less than 100,000. The greater variation in average wage found in smaller than in larger cities is a reflection of the relative impact of particular industries on cities of different sizes. The personal income of a one- or two-industry town will depend upon the wages paid out in those industries. In the larger city, with the greater dispersion of the labour force throughout a number of industries, wages in any one industry do not make as great an impact on average wage as they do in a smaller centre.

The town or small city blessed with a growth industry or with an industry which pays higher wages will tend to have a higher average wage level than other towns without this industrial advantage. The nature of the industry plays a major role in determining wages and income. Hanna found in his study of interstate wage differentials that variations in industrial composition accounted for 76 percent of observed differences in state earnings of wages and salary workers (1957, p. 151).

A confounding factor in the urban size-income association is the regional variations found in any large country such as Canada or the United States. The relationship between average wage of male wage earners in Canadian cities and the regional location of those cities was discussed in pages 214 and 215 of this thesis. In 1951 Ontario was the only region with more cities above the median wage than below it. All of the Quebec cities were below it. In 1961 only one city east of Ontario was above the median and the eleven lowest ranking cities were located in Quebec and the Maritimes.

Mansfield found similar regional disparities in his American cities (1957, pp. 290-294). Southern cities were consistently lower in average income than cities in other parts of the country. Among the smaller cities, income was higher in the Far West, while among the metropolitan areas, income was higher in the East North Central states.

It is fairly obvious that a city's income will be influenced by its regional location because part of the income of the region is the income of the city, and if the region is small or the city large, the income of the city may make up a large part of the regional income. The region *per se* does not influence the income of a city, but rather the fact that the city is located there places it under the same economic forces as the region itself. If the region is geographically separated from potential markets, the costs of transportation of goods will adversely affect industries in the city just as they do industries throughout the region. Certain industries are attracted to a region because of the resources, markets or amenities available. An industry will select for its location one of the urban areas in that region, unless the industry is an extractive one. When it is established in a particular urban area, other industries which have linkages with the original industry will select for their sites the same town or one in close proximity with that in which the original is situated. Thus the economy of any city is tied in with the economy of other cities in that region. It is in this way that urban incomes are a function of the region in which the city is located.

Few large cities have persistently low average incomes. The wage level may drop in a city for a time if one of its major industries suffers a depression, as occurred in Seattle a few years ago when the United States aircraft industry reduced production, and more recently in Detroit when demand for automobiles decreased, but these situations are usually temporary. The industry will usually rebound or it would not be a major industry. Minor industries may suffer recessions without seriously affecting the income levels of a large city.

Agglomeration economies contribute to the higher levels of income in larger cities. Firms realise higher profits because of the available services, linkages with other firms, and the available market for their products. These advantages tend to manifest themselves in the form of higher wages. However, smaller cities can also benefit from agglomeration economies. In this case, the agglomeration does not occur internally, but must take place over several cities which are located within convenient transporting and communicating distance of each other. A cluster of interdependent cities can have many of the advantages of one large metropolitan centre and avoid the diseconomies of pollution and congestion. Such a cluster would be expected to exhibit many of the benefits found in large cities including higher wage levels.

The situation in southern Ontario supports this hypothesis. A large number of smaller cities are clustered around the metropolitan centre of Toronto. These smaller cities are able to take advantage of the financial services provided by Toronto and to direct the energies of their labour forces to other industries. King found that one of the characteristics of these cities was a paucity of business and financial services which were available in Toronto (see p. 222 of this thesis). Numerous linkages occur between the cities rather than within cities.

Mansfield distinguished between neighbouring and isolated cities in an attempt to discern an income differential along this dimension (1957, pp. 297-299). He found a consistent tendency for the level of median income to be higher among neighboring than among isolated cities although this association was blurred by the variation of income within each category of city.

The advantages of geographical proximity to the incomes of cities are quite clear although other factors must also be considered. The economies of the cities must first be viable before they can benefit from their linkages. The fact that the three prairie provinces have no clusters of urban centres suggests a very real disadvantage to this region. The planning in these provinces seems to imply that no effort is being made to capitalize on the agglomeration economies that accrue from clusters of cities. More money and thought is put into opening remote outposts in the northern frontier than in building up the smaller towns around Edmonton, Calgary, Saskatoon, Regina, and Winnipeg.

The Patterns of Growth of Urban Areas

As a city grows it necessarily expands in space and social scientists have proposed different models of growth. The earliest and best known is the concentric circle or zonal hypothesis of Ernest W. Burgess (1929, pp. 114-123). The essence of the hypothesis is that as a city increases in population, it expands radially forming a series of concentric zones. In the centre of the city is Zone 1, the central business district. It comprises retail outlets, office and civic buildings, banks, hotels and entertainment centres.

Zone 2 circumscribes the central business district. It consists of an inner belt of factories and an outer belt of deteriorating residences. This zone is subject to encroachment by the central business district.

The third concentric zone is a zone of "workingmen's homes" while surrounding it is a "zone of better residences". The third zone is largely inhabited by second generation immigrants while Zone 4 contains the homes of the middle class, native-born Americans in single family residences or apartments. Finally an outer zone, "the commuter's zone", consists of a ring of encircling towns and hamlets.

Burgess stated that as the population grew, the inner zones would expand and replace contiguous outer zones. He recognized distorting factors such as natural hills, rivers, lakes, and administrative zoning of land, but argued that the concentric zones of his model would show through such barriers.

A second model of urban growth was proposed by Homer Hoyt and was called sector theory (1939). He found that in terms of rent levels, residential neighbourhoods of like quality tended to be located in pie-shaped sectors. Different types of residential areas tended to grow outwards along specific radii.

He also postulated some generalizations about the position of high and low rent sectors. High income families tend to locate on higher terrain, or along a lake, river, or ocean shore, and along the fastest existing transportation lines. Low income familes tend to live on the least desireable land close to railroad, industrial, and commercial areas.

A third model formulated by Chauncy Harris and Edward Ullman is called the multiple nuclei model (1945). They argued that the land use pattern of a city does not evolve around a single nucleus, but around several. Examples of such nuclei might be an original retail centre, a port, a railroad station or a recreation area. Thus the central business district may develop around the original retail area; the wholesale and light manufacturing district around a railroad terminal; the heavy industrial district at one edge of the city where large tracts of land and transportation facilties are available. Higher class residential districts are located on high, well-drained land, away from the noise and pollution of industrial areas while lower class residences are located between the high class neighborhoods and the nuisance areas.

Harris and Ullman suggested that the rise of separate nuclei and differentiated districts could be the result of four factors.

1. Certain activities require specialized facilities. Such activities would locate where the specialized facilities are available. Thus a retail district must have a centralized location where it is accessible to most of the population.
2. Certain activities cluster in the same location because they have horizontal linkages, i.e., they attract similar customers or utilize common facilities. Thus we often find several entertainment facilities in one part

of the city and several specialty retail shops on one street. Their locational proximity attracts customers who know they will have a wide selection of goods from the numerous outlets.

3. Certain unlike activities are adversely affected by their mutual proximity. Thus heavy industry requires large tracts of land and emits noise and air pollution while retail areas benefit from clean air and quiet environment, and must have a high density of units in order to be amenable to pedestrian traffic.
4. Some activities cannot afford the high rent required for location in choice areas. Thus lower class housing is located in less desireable areas because higher class families are not competing for those areas.

The three models, concentric zones, sectors, and multiple nuclei, are over-simplifications of the real world. Burgess and Harris and Ullman used land-use as their criterion while Hoyt used land rent as the exclusive criterion in his model.

Peterson criticizes Burgess for not including industry in his schema and for designating the use of the suburbs exclusively to upper-middle class residences (1975, p. 478). Critics have suggested that more adequate models would recognize other important factors in the shape of the city. Such factors would include the effects of administrative boundaries, the attachment of upper-income groups to historic and prestigious islands close to the urban core, the process of urban renewal, and the value-systems of the residents and of city planners.

The three models reflect the periods in time when each was conceived. Burgess' model was first publicized in 1929 before the automobile had made its impact, and cities were consequently highly centralized. Hoyt's model appeared ten years later. By this time the automobile had begun to have an effect and the process of urbanization was gaining momentum. Finally the Harris-Ullman model reflects the increased size of American cities and the resultant urban sprawl which prohibits a city of over one million from having a single centre or nucleus.

In spite of the inadequacies of each of the three models, they contribute to our understanding of the structure and growth of many North American cities. A number of factor analytic studies carried out since the Second World War indicate that the three models are "independent, additive contributors to the total socio-economic structuring of city neighborhoods" (Berry, 1971, p. 100).

Two of the studies were conducted in Canada, in Toronto (Murdie, 1969), and in Winnipeg (Nicholson and

Yeates, 1969). Both studies confirmed the results of similar studies in the United States. Household attributes and family status tend to vary radially as indicated by the Burgess model. Toward the core of the city buildings tend to be older and densities higher. More apartments, more rental accommodation, fewer children, and more elderly persons are located there.

Social rank and economic status tend to vary axially, as proposed by Hoyt. Lower income residences tend to grow out from the core of the city along transportation routes, industrial sectors and low terrain. High income areas extend axially from the city centre along higher ground, river banks or lake shores. Associated with high income are such attributes as occupation, education, value of home, and amenities available in the home.

The multiple nuclei model seems most appropriate when applied to ethnic and racial dimensions. If we compare numerous cities we find ethnic neighborhoods situated randomly in various parts of the cities. Associated with ethnicity are such attributes as country of birth, religion, citizenship, and period of immigration.

This section has dealt with three major models of internal urban structure, and consequently reflects the prevailing ideas on the patterns of growth of urban areas.

Burgess suggested the basic thought that as a city grows it expands outwards while retaining its internal structure. Thus each zone presses upon and eventually succeeds in dominating the contiguous zone that contains it. Hoyt recognized the same centrifugal growth forces but focused on the movement of specific sectors out from the core rather than viewing growth as the expansion of concentric zones in a balanced, geometric formation. Harris and Ullman adopted a radically different approach to urban growth, an approach which is most applicable as cities grow larger and more complex. Their model has a city growing as the result of the expansion of numerous nuclei scattered more or less randomly throughout the city.

Systems of Urban Density

Urban population density has been a much maligned phenomenon since Wirth's article in 1939. He suggested in that article that density encouraged an impersonality in relationships, a view of others as objects, and a tendency to exploit these objects. His article presented an unsubstantiated and one-sided view of the consequences of urban density. He ignored a number of advantages. Persons living in densely populated cities are offered a greater number and variety of alternate goals and means of attaining their goals.

They have access to far more styles of gratification, services, and selective personal association than their counterparts in rural areas.

Another advantage of density is that it is an economizing circumstance; it minimizes the time and cost of exchanges of goods and services (Hawley, 1972, p. 523). The cost per household of utility lines, street surfacing, sidewalk construction, police and fire protection, and delivery service is lower as density increases.

Large aggregates of people attract industry because they represent a large domestic market with guaranteed low costs of transporting goods to consumers. They also assure new industries of a large labour force pool and of the presence of other industries with the attendant linkages that reduce costs of production.

Another aspect of urban density, and the one that is of concern in this thesis, is the density and distribution of cities in a region or country. If industries and people derive benefits from their mutual proximity within a city, then they would also benefit from the proximity of a number of cities.

A mathematical relationship between the rank and size of a system of cities has been proposed by G. K. Zipf (1949). The relationship is as follows:

$$P_r = \frac{P_1}{r^q}$$

where P_r = the population of the rth ranking city,

P_1 = the population of the largest city,

r^q = the rank raised to an exponent q, which generally has a value very close to unity.

Zipf argues that a country whose cities conform to this relationship between city size and city rank has an integrated national system of cities.

The rank-size relationship bears on another concept, that of the primate city. Primacy occurs when one city in a system is several times larger than any of the others. It is an alternative to the rank-order relationship because it is applied to systems where a stratum of small towns and cities is dominated by a single larger city and where the distribution of cities varies from that prescribed by the rank-size rule (Berry and Horton, 1970, p. 66). A country whose urban system does not approach the rank-size regularity is classified as over-urbanized, a characteristic found most often in under-developed countries and said to be the result of superimposed colonial economies in these countries (Berry, 1964b,p. 139). Although the rank-size rule lacks any theoretical grounding, it is supported in varying degrees by empirical investigations.

One of the most comprehensive studies of the rank-size rule was carried out by Berry (1964b). He fitted curves to 38 city size distributions selected simply on the basis of availability of data. Thirteen of the 38 countries conformed to the rank-size rule, including countries as large as China, as small as Switzerland, as highly developed as the United States and as poorly developed as Korea.

Fifteen of the countries had primate city size distributions in which there was a gap between lesser city sizes and one or more large primate cities. All fifteen countries are small with highly specialized agricultural economies (Berry, 1964b,p. 141) and range from underdeveloped Thailand to Denmark and the Netherlands.

The remaining nine countries, including Canada, have city size distributions intermediate between the rank-size rule and the primate distribution.

In an attempt to make sense out of these classifications, Berry ran a chi-square test to determine whether a relationship existed between degree of urbanization and type of city size distribution. The degree of urbanization of countries was determined by the percentage of population in cities of 20,000 or more. The countries were placed in six categories of urbanization for convenience of analysis. The chi-square test revealed no relationship between the degree a country was urbanized and the type of city size distribution it had.

Berry also compared the city size distributions of the 38 countries with an alternate index of primacy, the ratio of the population of the largest city in a country to the aggregate population of the first four cities (Berry, 1964b,pp. 144-147). He found that countries which had once been dominated by colonial powers had primate cities which were national capitals as well as cultural and economic centres. Small countries which had once been large empires also had primate cities which had acted as empire capitals. Countries with more than one primate city, such as Canada, tended to be larger countries made up of two or more disparate regions or included specialized cities whose functions were complementary rather than duplicative.

In a further attempt to extract a theoretical framework from the empirical distribution of city size, Berry suggests that primacy occurs when few forces are affecting the urban system while the rank-size rule applies when the system has been subject to a wide variety of forces. The capital city in underdeveloped countries exists because of the former ties that country had with its "mother country" and because of the forces generated by the political and administrative spheres centered in that city. On the other hand, rank-size distributions are found in countries subject to a wide variety of economic, cultural and political forces over a long period of time.

Berry proposes several sub-hypotheses from this framework:

> fewer forces will affect the urban structure of a country, (a) the smaller is that country, (b) the shorter is the history of urbanization in the country, (c) the simpler is the economic and political life of the country and the lower its degree of economic development (Berry, 1964b, p. 149).

Canada ranked as an intermediate country, conforming neither to the primacy nor to the rank-size model. It tended toward the rank-size model but had more than one primate city. One inference drawn from the sub-hypotheses and applied to Canada is that the country is so large relative to its population that it requires regional primate cities. Another reason for its divergence from the rank-size model might be that Canada was a colony and has been dependent on England, and more recently the United States, for its economic and cultural development. A third reason might be the short span of time in which the Canadian urban system has existed. Perhaps, given more time, the system of cities will conform to the rank-size model.

Berry then investigated an hypothesis that rank-size distribution is associated with economic development.

He obtained indicators of economic development by factor analyzing a data matrix comprising the ranks of 95 countries on each of 43 variables considered to be indicators of economic development. He obtained four factors. They were a factor of technological development, a factor of demographic development, a low national income factor, and a high per capita-low per unit area factor. He then compared the types of city distributions with each of the four derived factors of economic development. His conclusion was that "different city size distributions are in no way related to the relative economic development of countries" (Berry, 1964b, p. 150).

A case study of Poland provides valuable insights into the complex processes linking urbanization, economic development, and the distribution of city sizes (Berry and Horton, 1970, pp. 75-87). Poland has shown a fairly rapid rate of urbanization since the Second World War, because of natural increase, rural-urban migration, and administrative changes of urban boundaries or creation of new towns.

Natural increase was responsible for 41 percent of total urban growth in Poland from 1950 to 1959. High natural growth in that period was due to the fall in the death-rate under conditions of improved sanitation as well as to the post-war baby boom. However economic factors were probably behind these demographic changes (Berry and Horton, 1970, p. 77). A rise in the living conditions brought about by the social revolution and later by rapid industrialization made a powerful contribution to the change in birth and death rates.

The rank-size rule gains empirical support in the distribution of city size in Poland. This suggests an ideally balanced structure of the urban network with a multitude of counter-balancing forces operating on an older, established urban system in a reasonably large country (Berry and Horton, 1970, p. 79).

Urbanization trends in Poland are fairly typical of countries with a tradition of urban life and passing through a period of rapid industrialization. However urban growth in Poland may be better balanced than in many other countries because of the limited growth of Warsaw and the development of numerous large and middle-sized towns. Any regional variations that occur are largely the result of variation in the distribution of natural resources, variation in the time of development of urban areas, differences in the functions and sizes of towns in different regions, and differences in the population structure of various states.

Canadian urbanization has tended to locate in clusters. Major concentrations of urban centres are located in the lower Great Lakes and the St. Lawrence Valley, the lower Fraser Valley, the Assiniboine and Red River Valleys, and the Atlantic and St. Lawrence Gulf coastlines. Only in Alberta and Saskatchewan are urban areas not clustered. These two provinces have strong agricultural traditions in which the urban settlements have functioned primarily as central places for the surrounding rural area. This may explain the greater distances between urban areas in these two provinces than in the others.

The prairies have been the least urbanized of the major regions in Canada although they surpassed the Maritimes during the 1950's. The Maritimes have consistently lost rank with respect to percentage of the population situated in urban areas since the eighteenth century when they may have been the most highly urbanized of the major regions (Stone, 1967, p. 37). Since 1961 they have ranked lowest while the prairie provinces have made rapid gains in urban growth as a consequence of highly profitable expansion of oil and gas production. Since 1941, they have commanded a considerable lead in rate of urbanization.

A comparison of the independent but contemporary development of the urban systems of Canada and Australia provides some insights into the pattern of growth of Canadian cities. Both countries have seen an extremely rapid growth of large metropolitan centres and a concomitant stagnation of smaller urban areas and rural regions (Bourne, 1974, p. 153). Both countries provide similar frameworks within which an urban system can evolve. Both are large geographic areas of recent settlement with resource-based economies dependent on external trade for further growth.

The pattern of urban development of both countries conforms with that of many colonized countries. The original urban points were ports having strong linkages with the mother country. As the exploitation of resources in the colony gained momentum, the urban system spread inland, but always retained close linkages with the original port or ports which monopolized trade, polarized cultural amenities, trade and administrative functions of the entire colony. The inland towns and cities played a supportive role within the total urban system, and acted primarily as service centres in the agrarian-dominated or staple-based economy. In many countries, including Canada and Australia, this urban system evolved into a more diverse and technologically integrated pattern of cities with the original primate cities retaining their primacy but losing the overwhelming dominance which they had held during early development.

There are striking differences between the urban development of the two countries in spite of the similarities they share as resource-based colonies providing raw materials to Great Britain. Australian settlement concentrated in coastal towns with little spread into the interior. Australia was an urban nation before it became a rural one (Bourne, 1974, p. 154). The urban background of the early settlers, shortage of navigable rivers, and inhospitable environment combined to discourage the spread of settlements inland. Canada, on the other hand, was blessed with one of the finest inland waterways in the world - the St. Lawrence River and Great Lakes water system. The early inhabitants were motivated to utilize this transportation system by the wealth of furs, and later, minerals, which vied with the bountiful fisheries off the Eastern seaboard as major propulsive industries.

Another distinguishing factor in the urban growth of the two countries is the isolation of Australia as opposed to the proximity of the United States to Canadian cities. The Australian pattern of settlement was independent of any other economic and demographic system. In contrast, Canadian cities developed in spite of, and often because of, the more advanced urban system to the south. The clustering of urban areas in southern Ontario is primarily due to their proximity to large markets and head offices in the industrial belt of the north-eastern United States.

As the Canadian population increased, settlement spread across the southern strip of Canada in accordance with the transportation system, available resources, and the ties and conflicts with the United States urban system. The bulk of the Australian settlements remained on the coast, providing ready access to over-seas trade centres. Australian urban growth appears to stem from a highly localized multiplier effect whereby investment in a city resulted in growth and further investment in that particular city, rather than from the inter-urban linkages that characterize Canadian urban growth (Bourne, 1974, p. 155).

The growth of these settlements which precluded the spread of population into the hinterland is reflected in the very high proportion of Australian population now living in urban areas. In 1971, 83 percent of the total population, 12.7 million, was classified as urban, while 76 percent of Canada's 21.6 million was so classified. The percentages are comparable because the two countries use approximately the same definitions of urban and rural. In Canada the urban population is found in those incorporated urban areas of 1,000 or more population, while Australia requires a cluster of 1,000 or more population, and 750 or more in Tasmania.

The Canadian proportion living in urban places approximates the average level of urbanization in western industrialized nations, while Australia exceeds this level to be one of the most highly urbanized countries in the world. It achieves this in spite of its large land area and the development of its resource industries. Such a high level of urbanization reflects the historical settlement patterns of Australia with heavy concentration in the cities and corresponding sparse settlement in the hinterland.

The spread of settlements into western Canada resulted in rapid growth of western cities which were settled later than their eastern counterparts. It led to an instability of growth in the urban system as younger cities surpassed older, less viable centres in size and economic dominance. The population of Quebec was surpassed first by Montreal, then by Toronto, Ottawa, Hamilton, and finally Vancouver, Winnipeg, Edmonton, and Calgary. As the urban hinterland of Toronto spread west to the Pacific, its population approached that of Montreal which looked to French-speaking Canada for many of its cultural and economic linkages. The growth of Australian cities was much more stable, with each holding its rank, or making a minor shift, throughout Australia's history (Bourne, 1974, p. 158).

The application of the rank-size rule to the systems of cities in the two countries offers interesting equivalence and contrast. Neither country has a system of cities that conforms to the rule, although neither diverges from it to a considerable extent. Canada's system of cities conforms to the rule more closely than does Australia's. Both countries show a degree of primacy, and in each case the primacy lies in two cities rather than one: Montreal and Toronto in Canada, Sydney and Melbourne in Australia. Australia shows a considerable gap in middle-size cities. The fifth and sixth ranking cities in 1971 are Perth (640,000) and Newcastle (250,000). Canada has nine cities falling in this population range. This gap in Australia's rank-size distribution is a reflection of its smaller population and the apparent lack of linkages which work to balance economic growth within a system of interconnected cities.

Bourne pursues an interesting parallel, although perhaps not a very relevant one, between the geographical distributions of the urban systems in the two countries (1974, pp. 165-168). He notes that the cores of the urban systems lie in the south-east of each country, and that the urban network stretches north-east from that core for a relatively short distance and west across the southern boundary of each country to the west coast. In each country, the core consists of two large industrial centres, Toronto-Montreal and Sydney-Melbourne, with the national capital placed strategically between them. Thus Brisbane seems to hold the same network position relative to Sydney as Quebec does relative to Montreal. Adelaide and Winnipeg hold similar positions in the network serving as links to the west. The similarities in the positions of Perth and Vancouver as the western extremities of the urban systems is obvious.

Bourne raises the question whether these parallels have significance (1974, p. 168). The geographical similarities are probably trivial but what may be of interest to students of urban development is the effect of the geographical position of each city within the urban network on the growth of that city. Both countries presently experience rapid urbanization in the core region and slower growth or stagnation at other points along the urban network. High densities of population at the core are contributing to wage inequalities, traffic congestion and inadequate services for large, densely populated regions.

The Age-Sex Distribution

The age-sex distribution reflects historical changes in demographic rates such as fertility, mortality, and migration which may be associated with social and economic factors. The age-sex distribution plays an important role in determining what services local governments must provide in an urban area. A large number of pre-schoolers necessitates the allocation of government funds to education facilities while a large number of older persons requires considerable government investment in health services and residences for the retired.

Stone characterizes the urban population of Canada as "female-dominant" (1967, p. 44). Females outnumbered males in the total urban population in the age group 20 - 34, and in the older age brackets because of their lower mortality rate. The urban masculinity ratio for all ages was 98.2 compared to the rural masculinity ratio of 112.2 (Stone, 1967, p. 46).

The urban population was also older than the rural population in 1961. The median age of urban males was 27 while that of rural males was 23. The median age of urban females was 28 while that of rural females was 22 (Stone, 1967, p. 47). These differentials reflect an urban population considerably older than the rural population. Lower fertility rates in the urban areas can account for this, as well as rural-urban migration, which occurs predominantly among persons 20 years and older. Retiring couples from farms often move into urban areas, contributing to the higher median age, but on few occasions do retiring urban couples move to rural areas.

The age-sex structure of Canadian urban areas in Canada varies according to the size of the urban place. A definite tendency occurs for the age-sex structure of smaller urban settlements to emulate the rural age-sex structure. The masculinity ratio increases as urban size declines. Masculinity ratios range from 97.6 in cities of 100,000 and over to 100.4 in urban areas of 5,000 and less in 1961. Median age of males decreases from 28.1 in cities of 100,000 and over to 24.8 in urban centres of less than 5,000, while corresponding values for females are 29.2 and 25.3 (Stone, 1967, p. 53).

Some interesting regional differentials occur among Canadian cities in 1961. The level of female domination decreases as one moves westward until British Columbia where the masculinity ratio is 99.8 (Stone, 1967, p. 54). The median ages of urban populations are lowest in eastern Canada. The median age in urban centres in the Maritimes is 25, in Quebec, 26, while the regions with the highest median ages are Ontario, 29, and British Columbia, 31.

Another aspect of urban age-sex structure by region is the age-sex pyramid of the Maritimes which most closely approximates a normal pyramid characterized by declining percentages in each five-year age group from 0-4 to 85 and over. This reflects the inability of Maritime urban centres to attract and hold migrants. The other four regions have age-sex pyramids which bulge at age groups 25 to 44 indicating a certain amount of in-migration and much less out-migration of persons who have reached labour force participation age. The largest bulge occurred in the Ontario pyramid.

Ontario and British Columbia have urban populations with the highest median ages. British Columbia urban settlements attract retiring couples, raising the median age by increasing its old-age dependency ratio and by maintaining a low youth dependency ratio.

The old age dependency ratio is calculated as follows:

$$0 = \frac{P_{65+}}{P_{15-64}} \cdot 100$$

where 0 = old age dependency ratio,

P_{65+} = population 65 years and over,

P_{15-64} = population ages 15 to 64.

The youth dependency ratio is calculated in the same way except that P_{65+} is replaced by P_{0-14} which is the population less than 15 years of age. The distribution of urban youth dependency ratios and urban old age dependency ratios by region in 1961 is as follows (Stone, 1967, p. 54):

	Urban youth dependency ratio	Urban old age dependency ratio
Maritimes	58.7	13.4
Quebec	54.7	9.5
Ontario	51.2	13.3
Prairies	54.9	14.2
B.C.	49.3	18.6

Ontario achieves a high median age not by having a high old age dependency ratio (three regions have higher ratios) but by having a low youth dependency ratio. Only British Columbia's youth dependency ratio is lower.

Quebec's age-sex pyramid is notable for its low old age dependency ratio while the Maritimes pyramid shows a large youth dependency ratio. These two separate factors contribute to make Quebec and the Maritimes the two regions with the youngest urban populations.

Two aspects of the urban age-sex structure appear to be associated with economic standards of the five regions. If regions are ranked according to economic standard, the median age of urban populations decreases with decreasing economic standard, with one exception. British Columbia urban areas have a higher median age than Ontario although Ontario's economic standing is higher than that of British Columbia. The masculinity ratio is also associated with economic standing with one exception. Urban areas in regions with highest economic development are least dominated by females, with the exception of Ontario which ranks third instead of first:

	Economic Rank	Median Age Rank	Masculinity Ratio Rank
Ontario	1	2	3
B.C.	2	1	1
Prairies	3	3	2
Quebec	4	4	4
Maritimes	5	5	5

A definite causal relationship can be established between economic rank and median age rank. Lower fertility rates are normally associated with higher economic development. Low fertility contributes to a higher median age. Consequently we would expect median age to be higher in populations that are more developed economically. However this explanation loses credibility when the youth dependency ratio, which is a more direct indicator of fertility is found to have a weaker association with economic rank than does median age rank.

	Economic Rank	Urban Youth Dependency Rank
Ontario	1	4
B.C.	2	5
Prairies	3	2
Quebec	4	3
Maritimes	5	1

The masculinity ratio is more difficult to tie in with economic development. Industrialization and urbanization have provided many more opportunities for women in the labour force than exists in primary economic activities such as agricultural and extractive industries. Certain industries, such as service industries, tend to attract more females than do other sectors of the economy. However services, and especially female dominated services, are not necessarily associated with higher economic development. One might hypothesize that female domination of urban population is associated with standard of regional economic development on the basis of rural-urban differentials in the masculinity ratio, but this would not be supported by the data. To make the opposite hypothesis, that economic development is associated with higher masculinity ratios, has no theoretical support even though it describes the 1961 Canadian data. The striking aspect of the masculinity ratio is that it increases from east to west. This suggests that female domination is associated more closely to types of industry than it is to level of economic development.

The urban population in Canada increased by 52 percent over the 1951-1961 decade. This growth can be broken down into several components:

1. increase in population within existing urban boundaries;
2. change in classification of an area from rural to urban minus change in classification of an area from urban to rural;

3. population increase in an area which was reclassified from rural to urban.

In the decade 1951 to 1961, more than 75 percent of the increase in urban population was due to population increase while less than 25 percent was due to reclassification of the area (Stone, 1967, p. 92). Population increase in turn can be decomposed into natural increase and net migration. Net migration was responsible for 43 percent of total population increase in urban areas (Stone, 1967, p. 95).

Net migration does not reflect total population movement, which is out-and in-migration, but it does reflect the attractiveness of the area to in-migrants and its holding power to potential out-migrants, relative to other urban areas in Canada. Net migration to urban centres from 1951 to 1961 amounted to 16 percent of the 1951 population of those centres (Stone, 1967, p. 99). Urban areas in the Maritimes experienced zero net migration while centres in the Prairies were increased by 28 percent. Contributions of net migration to the remaining three regions were relatively close to the national average.

The growth of larger urban centres was due to a proportionately greater extent, to net migration than was the growth of smaller centres. Net migration in cities of

100,000 or more accounted for growth that was 23 percent of the 1951 population. In cities of 30,000 to 100,000, net migration contributed to 17 percent of the population increase while in smaller urban complexes it accounted for only about a four percent increase over the 1951 population (Stone, 1967, p. 95).

Natural increase played a larger role than net migration in population growth of urban areas. Predominance of natural increase holds true for all size categories except the largest. In cities of 100,000 or more, natural increase accounted for 44% of total population growth.

The increase due to natural increase in all other size categories is more than 50 percent of total population growth. Large cities have a disproportionate attraction to migrants and this accounts for the differential growth rates by city size. Large cities have higher growth rates in spite of the fact that natural increase has a relatively constant effect on population growth rate for all city size categories (Stone, 1967, pp. 96-97).

Migration has an effect on the age-sex structure of a population which is very different from the effect of natural increase. It is age- and sex-selective although the age-categories and sex in which it is selective will vary over time and place. Migration to the Prairies at the turn of the century was predominantly masculine because the occupations made available were for males. In cities with expanding service facilities, migration may be predominantly female. Age selectivity in British Columbia, where older people retire, is different from that of Alberta where most in-migrants are seeking jobs.

Data on age and sex differentials of migration in Canadian urban areas for 1951-1961 must be collected indirectly because such data are not available. Stone was able to make calculations for a few of the larger cities, using the life table survival ratio technique. This method involves the application of age-specific life table survival rates to a population resident in an urban area at the beginning of a time period to determine how many in each age bracket would survive to the end of the time period. The residual of the estimated survivors and the population as measured by some other means, usually a census, is a measure of net migration. The problem of estimating net migration of the youngest age group(s) remains because persons in this group were not alive at the beginning of the period. One method of estimating net migration for this age group is to apply child-woman ratios of the observed end-of-period population to the estimated survivors and obtain a residual from the derived number of children and the number in the population.

Stone applied this technique to a group of larger Canadian cities which included Halifax, Saint John, Montreal, Quebec, Kitchener, London, Ottawa, Sudbury, Toronto, Windsor, Calgary, Edmonton, Regina, Saskatoon, Winnipeg, Vancouver, and Victoria (1967, p. 102). He found that male net migration ratios were higher than female net migration ratios in 1951-1961. A net migration ratio for one sex is the ratio of net migration of that sex and the 1951 population of that sex. Females tend to dominate specific migrant age groups, namely ages 10 to 24 and 60 plus while males tend to dominate all other ages (Stone, 1967, p. 105). Females show the greatest dominance in ages 15 to 24 while male migrants are most dominant in ages 30 to 39.

When the distribution of age of migrants is considered separately for each sex, females are seen to rise gradually from age 10, steeply from age 15 and to peak at about age 25, after which it drops rapidly to age 40. The male distribution rises less rapidly, peaking at about age 30, after which it declines less rapidly than does the females distribution.

Some striking regional differences appear in the distribution of net migration ratio by age. Only British Columbia and the Prairies have a considerable amount of positive net migration for ages 35 and more while the net migration ratio for these ages in the Maritimes is negative. The largest positive ratios occurs in the Prairies for both males and females in their twenties. The Quebec and Ontario age profiles are quite similar to those for the entire set of cities in the sample. The Prairies are notable for their large positive net migration ratios for persons under 35, British Columbia for its positive ratios for persons over 35, and the Maritimes for their low positive values for persons under 35 and for large negative values for persons over 35 (Stone, 1967, pp. 108-109).

It was noted in previous paragraphs that the essential differential of growth between larger and smaller urban centres is net migration. In-migration accounts for the more rapid rate of growth of larger cities. It also has the effect of distorting the age-sex structure of the city because it is age- and sex-selective. Generally, its effect on the urban age-sex structure is to reduce the masculinity ratio and create a bulge in the pyramid at the ages of labour force entry and the early years of working life.

Stone constructed hypothetical populations for six cities in Canada on the basis of their populations being closed to migration since 1901 (1967, pp. 113-117). He found that migration had the effect of lowering the masculinity ratio from 96.9 in the hypothetical population to 95.0 in the actual population in 1961. Migration lowered the median age of the population below that of the hypothetical, closed population. The median age of males, which was 29.7 in 1961, would have been 34.5 under conditions of no migration, while the corresponding values for females were 29.7 in the actual population and 36.5 in the hypothetical population. The "younging" effect that migration has on the population is also reflected in the increased youth dependency ratio, 41.6 compared with a ratio of 38.4 in the hypothetical population, and the reduced old age dependency ratio, 13.0 in the real population, compared with 18.7 in the closed population. Migrants have this effect because they are predominantly younger persons who will have children after becoming established in their new habitat. Their children contribute to a higher youth dependency ratio while the migrants themselves reduce the old age dependency ratio by enlarging the denominator of this ratio.

An analysis of the age-sex structure of Canadian Metropolitan Areas in 1961 provides some insights into the effects of differential urban in-migration. Table 31 displays masculinity ratios, median ages for each sex, and dependency ratios for youth and old age. Masculinity ratios range from 93 in Quebec to 107 in Sudbury. When these two extreme values

Table 31

Age-Sex Distribution Statistics of Population of Census Metropolitan Areas, 1961

	Masculinity Ratio	Median Age		Dependency Ratio	
		Male	Female	Youth	Old Age
Calgary	101	26.8	26.7	57	12
Edmonton	102	25.8	25.2	59	10
Halifax	101	24.3	25.4	54	9
Hamilton	99	28.5	29.4	53	13
Kitchener	98	27.5	28.8	53	13
London	96	28.4	29.6	50	14
Montreal	97	27.3	28.5	49	10
Ottawa	96	25.3	27.3	56	11
Quebec	93	24.8	27.1	51	10
Saint John	97	26.3	28.4	56	15
St. John's	95	22.2	22.8	63	11
Sudbury	107	24.1	22.4	63	6
Toronto	98	29.8	30.8	46	12
Vancouver	99	31.4	32.2	48	18
Victoria	97	31.4	35.6	49	28
Windsor	99	28.1	28.9	56	14
Winnipeg	97	28.7	29.6	49	15

SOURCE: Adapted from Table 9.3, Stone, 1967.

are excluded, the range of masculinity ratios is rapidly diminished to include values from 95 in St. John's to 102 in Edmonton.

Median age shows considerable variation among the M.A.'s. Male median age varies from 31.4 in Vancouver and Victoria to 22.2 in St. John's. Female median age varies from 35.6 in Victoria to 22.4 in Sudbury. The youth dependency ratio tends to vary inversely with median age of each sex and with the old age dependency ratio, the latter three being closely associated. Thus, the highest youth dependency ratio, 63, is found in St. John's and Sudbury, where the lowest median ages of males and females occur. The highest old age dependency ratio, 28, occurs in Victoria where the highest median age for both males and females is found.

The "oldest" age-sex structures are found in Vancouver, Victoria and Toronto while the "youngest" are in St. John's, Sudbury and Halifax. Other cities that approach these two extremes include cities with "older" populations such as Hamilton, Winnipeg, London, Windsor, and Montreal, while cities with younger populations include Calgary, Edmonton, Ottawa, Quebec, and Saint John.

The local age-sex structure is influenced by the history of local fertility which is strongly affected by urbanization. Place of residence has been ranked third behind religious affiliation and social class as one of the most powerful influences on fertility (Thomlinson, 1969, p. 97). One reason for this may be that the urban environment provides alternative work opportunities and recreational opportunities which compete with children for women's attention. Another explanation may be that children cannot make the same economic contributions to the family in town that they can on the farm where they are capable of caring for livestock and gardens. A third possible explanation is that living in closer proximity to other women, a young married female may be exposed to knowledge of birth control practices of which she would never learn if living in the country. Whatever the reason, fertility rates are lower in urban than in rural areas. In a recent study in Ontario, the total fertility rate, which is the sum of age specific fertility rates, was calculated for each county from 1961 to 1969. The overall nine-year weighted average for counties predominantly rural was 3496 while the average for predominantly urban counties was 2946 (Barratt, 1971, p. 11).

This difference in fertility by residence has been present in Canada since the middle of the nineteenth century (Henripin, 1972, p. 79). It is differentiated not only by rural-urban contrasts but it tends to be distinguishable in terms of city size. Birth rates are lowest in metropolitan centres and highest in small towns (Thomlinson, 1969, p. 97). Henripin measured the fertility of Protestant, English-speaking women aged 45-49, born in Canada, with a high-school education, whose husbands have a secondary education and earn between $3,000 and $5,000 yearly. After controlling for all of these variables, he found that in relation to the overall average for this group of women, to which he gave a fertility index of 100, the fertility index for women with the same characteristics but in different residences was as follows: metropolitan areas, 88; urban areas of 30,000 to 100,000, 97; urban areas of 5,000 to 30,000, 105; urban areas of 1,000 to 5,000, 115; rural non-farm areas, 126 (Henripin, 1972, p. 115).

There is a considerable variation in fertility rates in the larger urban centres of Canada.

> The fertility of women between 20 and 35 years in St. John's is 50% higher then that of women of the same age in Toronto; that of women over 45 years is 100% higher in St. John's than in Toronto and, still compared to Toronto women, the fertility of women in Quebec City, aged over 55, was even higher in 1961 (Henripin, 1972, p. 91).

The fact that urban-rural fertility differentials are diminishing implies that it is not the residence per se that is influencing fertility so much as a number of other factors associated with residence. Some of these might include education, income, religion, husband's occupation, and wife's labour force status. However the persistent inverse relationship between urban size and fertility rate suggests that the urban size must be retained in any causal model of fertility.

Some Studies of Demographic and Urban Change

A number of studies have been carried out which investigate differences among towns and cities and the demographic differences of these urban areas shall be the focus of the remainder of this section. Moser and Scott analyzed the social, economic, and demographic variation of 157 local authority areas in England and Wales having urban populations in excess of 50,000 (1961). Included in the data they used were population size and structure, population change, households and housing, economic character, social class, voting, health, and education.

Moser and Scott placed a strong emphasis on demographic dimensions of urban differences. They extracted four major components from 57 variables using a common factor analytic model. The component accounting for the largest percentage of total variance was one they called "social class". The remaining three of the four largest components were demographic. The second largest component, accounting for 13.2 percent of the total variance was called "population change between 1931 and 1951"; the third component, accounting for 9.8 percent of total variance, was called "population change between 1951 and 1958", while the fourth component, accounting for 7.3 percent of total variance, was called "overcrowding" (Moser and Scott, 1961, p. 14).

The results of factor analysis are entirely a function of what variables have been selected to be factor analyzed, and no factor analytic model can be properly evaluated without knowledge of the input variables. Table 32 lists the 57 variables selected in this study. No theoretical justification is given for selecting these variables over others. In spite of this weak theoretical grounding, we can be certain that in Moser's capable hands there are no statistical oversights in the study.

Moser and Scott arbitrarily took product-moment correlation coefficients of 0.50 or better, and -0.50 or less, as being "substantial correlations" (1961, p. 60). No variables correlated substantially with population size. This implies that none of the variables listed in Table 35 are associated with city size. The weakness of size as a differentiating factor was also noted by Duncan and Reiss for American cities and towns (1956, pp. 130-133).

The one variable that did correlate most highly with urban size was percentage of the labour force in insurance, banking, and finance (r = 0.437). This coefficient obtains because such economic activities are concentrated in the largest cities rather than because of a linear association

Table 32

List of Fifty-Seven Variables Factor Analysed by Moser and Scott

Variable

- POPULATION SIZE AND STRUCTURE
  - Population ('000)
  - % of population aged 0-14
  - % of population aged 15-64
  - % of population aged 65 or over
  - Females per 1000 males
  - Females per 1000 males, aged 25-44
  - % of females aged 20-24 ever married
- POPULATION CHANGE
  - 1931-51 - total (%)
  - 1931-51 - births and deaths (%)
  - 1931-51 - balance of change (%)
  - 1951-58 (%)
  - Birth rate ratio, 1950-52
  - Birth rate ratio, 1955-57
  - % illegitimate births, 1950-52
  - % illegitimate births, 1955-57
- HOUSEHOLDS AND HOUSING
  - % persons in private h/hlds
  - % one-person h/hlds
  - % six-or more person h/hlds
  - % 1-3-room dwellings
  - Persons per room
  - % overcrowded h/hlds (composite)
  - % h/hlds at over 1½ persons per room
  - % h/hlds in shared dwellings
  - % h/hlds with piped water
  - % h/hlds with W.C.
  - % h/hlds with 5 amenities
  - % h/hlds with all 5, or only bath missing
  - New-housing rate, 1945-58--total
  - New-housing rate, 1945-58--L.A.
  - % L.A. of total houses 1945-58

Variable

- ECONOMIC CHARACTER
  - Occupied as % of total population
  - % of women in labour force
  - % in manufacture etc.
  - % in all service industries
  - % in retail
  - % in finance, etc.
  - % in professional services
  - Job ratio
  - Commuting ratio
  - Per capita retail sales, 1950 (£)
- SOCIAL CLASS
  - % in social classes I + II
  - % in social classes IV + V
  - Social class index
  - Social Survey J-index, 1954
- VOTING
  - General elections--
    - 1951 poll
    - 1955 poll
    - 1951 % voting left
    - 1955 % voting left
  - Local elections, 1956-58--
    - Voting in contested elections
    - % uncontested seats
    - Votes as % of total electorate
- HEALTH
  - Infant mortality rate, 1950-52
  - Infant mortality rate, 1955-57
  - Expectation of life at yr. One, 1950-52
  - T.B. notification rate, 1957
  - Mortality rate, 1957--lung cancer
  - Mortality rate--other cancer
  - Mortality rate, 1957--bronchitis
- EDUCATION
  - % with terminal edn. age under 15
  - % aged 15-24 in full-time education

SOURCE: Table 26, Moser and Scott, 1961.

between urban size and percentage of the labour force engaged in such activities.

Factor analysis provides a simplification of the 57 X 57 data matrix of product-moment correlation coefficients. Each extracted component from the factor analysis can be interpreted as representing those variables most highly correlated with it. The ten variables showing the highest correlations with the component called "social class" were (Moser and Scott, 1961, p. 71):

% in social classes I + II	0.931
Social class index	0.903
% aged 15-24 in full-time education	0.891
% in social classes IV + V	-0.871
% voting left, 1951 general election	-0.860
% voting left, 1955 general election	-0.859
% with terminal education under age 15	-0.845
sex ratio, 25-44	0.836
birth rate ratio, 1950-1952	-0.765

We note that this component is aptly called "social class". Two of the variables reflect attitude towards education, while three reflect social class structure of the population. Two of the variables are demographic. The sex ratio for ages 25 to 44 which is a femininity ratio, indicates that towns with high social class structures have higher proportions of females in the age brackets 25 to 44.

The sex ratio in these age categories reflects sex-selectivity of migrants. The high correlation of femininity ratio for ages 25 to 44 and the social class component indicates that a high proportion of females migrate to cities of a higher social class structure while males dominate the migrants to cities of lower social class structure.

The other demographic variable associated with the social class component is the average of the ratios of the local birth rate standardized to the national rate for 1950-1952. The association of fertility rate with social class is well known and does not require further elaboration here.

The second component is demographic. The ten variables showing the highest correlations with this component are (Moser and Scott, 1961, p. 72):

% of one-person households	0.817
Population change, 1931-51, due to natural increase	-0.773
% illegitimate births, 1950-1952	0.732
% illegitimate births, 1955-1957	0.688
Population change, 1931-1951 - total	-0.667
Per capita retail sales, 1950	0.643
Population change, 1931-1951, not due to natural growth	-0.630
% population aged 65 or over	0.628
% households with all 5 household amenities	-0.563
Commuting ratio	-0.557

Both natural increase and population increase not due to natural increase are negatively correlated with this component called "population change between 1931 and 1951". This suggests that those towns experiencing rapid increase due to natural increase also experienced large positive net migration. One must be careful when making this statement because the second variable, although measuring net migration, is also measuring another component of urban growth, annexation. One must also be aware of the possibility that although the two variables are negatively correlated with a third variable, they need not be positively correlated with each other. They are in this case, however; their correlation coefficient is 0.880. Ignoring the confounding effects of annexation, we note that this study reveals that cities with high rates of natural increase also increased because of high rates of net in-migration. Duncan and Reiss found the same relationship in American towns (1956, p. 210).

The correlation coefficients of the variables dealing with population change are negatively correlated with the component called population change, 1931-1951, indicating that the component is positively correlated with variables inversely related to population growth for that period. Illegitimate birth rates and per capita retail sales are highest in those towns experiencing slowest growth. Cities with larger percentages of their population over the age of 65 also tend to be slow growing, as do towns containing higher proportions of one-person households. This is reasonable in light of the fact that towns experiencing rapid growth do not have the housing facilities to provide as many one-person households as towns with static population growth. It is also reasonable that slower growing towns have larger proportions of their population in the older age categories because younger persons tend to migrate to those towns that are growing faster, leaving an older population in towns of slower growth.

The higher rate of illegitimacy in slower growing towns is more difficult to reconcile with standard concepts of urban economic development. Possibly the younger persons who do not migrate to find jobs get involved in other non-economic activities. The fact that per capita retail sales tend to be higher in slower growing towns is also difficult to interpret in the context of the standard belief that economic and population growth are complementary.

The variables that correlate negatively with the component, and therefore positively with urban growth, are percentage of households with piped water, W.C., kitchen sink, cooking stove, and fixed tub, and commuting ratio, which is the number of persons either living in the area and working outside it or working in the area and living outside it, per 100 resident occupied population, 1951. The association of households containing more amenities with urban growth can be explained by the fact that growing communities have newer homes, which in turn are built with these amenities. The incidence of higher commuting ratios in growing towns is more difficult to explain. Perhaps growing communities are characterized by some instability, with persons moving in and settling in areas distant from their place of work because they could not find a more convenient place of residence. Another explanation is that the growing town is expanding over its boundary, so that the high commuting ratio is an artifact of the inappropriate positioning of the boundary.

The third component is also demographic. Labelled "population change between 1951 and 1958", it correlates highly with the following ten variables (Moser and Scott, 1961, p. 72):

New housing rate, 1945-1958 - total building	0.776
Population change, 1951-1958	0.661
New housing rate, 1945-1958 - local authority building	0.636
% of women in the labour force	-0.610
% of the population aged 15-64	-0.587
Birth rate ratio, 1955-1957	0.584

% of the population employed	-0.546
Commuting ratio	-0.525
Households in shared dwellings	-0.467
Birth rate ratio, 1950-1952	0.456

This component is positively correlated with population growth from 1951 to 1958. A number of other variables are also positively correlated with the component, which hold no surprises to the analyst. New housing and high birth rates can be expected to be associated with rapid population growth.

Several variables are associated with population growth from 1951 to 1958 in ways that seem to contradict popular ideas about population growth and one conflicts with the association found with the second component, labelled "population growth from 1931 to 1951". Commuting ratio associates directly with growth from 1931 to 1951 while it associates inversely with growth from 1951 to 1958. This suggests that commuting ratio has a random association with population growth. The only other possible explanation is that conditions changed from 1931-1951 to 1951-1958 and that the opposing effects of variation in commuting ratio reflect the opposing effects of these conditions. The economic conditions from 1931 to 1951 were extremely diverse, with a serious depression occurring during the 1930's followed by wartime conditions followed by the post-war economic boom. Economic conditions during the 1950's were characterized by economic expansion in the early Fifties, followed by gradual decline to the minor depression of the early Sixties.

Three variables associate with population growth, 1951-1958 in ways opposite to that which standard belief about urban growth would have. The first is the percentage of the labour force composed of women, which decreases as population change increases. Larger cities tend to have larger percentages of females in the labour force because they provide more sales, office, and service jobs than do smaller towns. Job opportunities available in the immediate environment are one of the important factors determining extent of female labour force participation and larger centres usually offer more such opportunities than do smaller centres (Sweet, 1970, p. 195). Females in larger cities tend to be better educated and to have fewer children, both of which contribute to the liklihood that they will be employed. Larger cities also attract proportionately more single girls looking for jobs; these migrants increase the female labour force participation rate in larger cities. Finally, one tends to associate female labour force participation with urban size because of the concomitant increases of both these variables over the past few decades. However, the association is negative in England and Wales from 1951 to 1958 as well as from 1931 to 1951 (r = -0.089). The probable explanation for this is that sectors of the economy showing most rapid growth in England and Wales during this period were those with low female participation rates. When technical, engineering, and construction industries expand, the female proportion of the labour force decreases because such industries are male dominant.

The second variable which has a surprisingly negative association with population growth from 1951 to 1958 is the percentage of the population aged 15 to 64. In-migration is usually associated with rapid population increase because natural increase cannot account for all of the increase. Those cities that grow fastest tend to be the ones with largest rates of in-migration, which in turn contributes to a swell in the population pyramid in the working age groups of the population pyramid. The reverse condition characterizes the most rapidly growing cities in England and Wales from 1951 to 1958; these cities have proportionately fewer persons in the working age groups. The source of their growth appears to be their high birth rates. The ratio of local to national birth rates for 1955-1957 and 1950-1952 tend to be higher in those urban areas showing most rapid population growth from 1951 to 1958. This suggests that internal migration in England and Wales was at a low ebb if natural increase was the prime factor in growth of the expanding cities.

The third variable with an unorthodox relationship with urban growth is the employment rate. Cities with lowest employment rates are the ones growing the fastest. Theoretically, cities expand because they provide job opportunities which attract migrants. This results in high in-migration and high employment rates. Of course, occasions can arise where in-migration exceeds the rate of expansion of job opportunities in which case higher levels of unemployment can ensue. However, in cases of expansion, the market can adjust itself, and usually, expanding cities have lower unemployment rates than stagnating cities from which the migrants came. The association between low employment rates and faster growth rates in England and Wales from 1951 to 1958 is better understood in the context of the previous variable discussed. Migration played a minor role in urban growth, so the usual association with employment rate and urban growth, does not hold. Cities grew because they had higher birth rates, which have no direct influence on employment rates. The increments to the population pyramid appeared at its base, affecting the younger age groups. The population of working age is proportionately smaller in the faster growing cities, implying that full employment should prevail, because proportionately fewer persons per urban size are vying for jobs. In fact, the opposite occurs: employment rates are lower in these cities. One gets the impression that these cities exemplify the condition Richardson points out when he states that urban population growth and urban economic development can be conflicting concepts (Richardson, 1971, p. 80). The most rapidly growing cities in England and Wales from 1951 to 1958 are those with the highest birth rates, lowest employment rates, and lowest proportion of the population of working age. Unfortunately Moser and Scott do not provide an indicator of economic development with which to compare the 1951-1958 growth rate, so there is no way of investigating this inference.

The fourth component is called "overcrowding" and is correlated most strongly with the following ten variables (Moser and Scott, 1961, p. 73):

Households at over 1½ persons per room	0.608
Overcrowded households (composite index)	0.599
% poll, 1955 general election	-0.586
% poll, 1951 general election	-0.573
Persons per room	0.442
% 1-3 room dwellings	0.413

% households with piped water	-0.405
% working in manufacturing, agriculture, and mining	-0.389
% six- or more person households	0.375
% females aged 20-24 ever married	-0.366

Three of the variables relate to crowding while one measures percentage of dwellings that are small - less than four rooms. A fourth reflects large households, a fifth, small dwellings, and a sixth, presence of piped water in dwellings. This last variable would be expected to appear in the cluster of social class variables, but it does not. It correlates with this component, which is clearly a density component, and which is independent of the variables correlating with the social class component. Two other variables relate to voting behaviour: one to the industrial classification of the population and one to the marital status of females, 20-24.

In summary, Moser and Scott provide us with a comprehensive study of the demographic, social, and economic dimensions of towns in England and Wales from 1931 to 1958. They use the technique of component analysis to investigate how much of the total variability of the 57 variables can be accounted for in four new constructed components. The four components were named "social class", "population change, between 1931 and 1951", "population change, between 1951 and 1958", and "crowding". The two growth components are of interest here because of the usually close association between urban growth and urban development. The demographic variables associated with growth from 1931 to 1951 are lower illegitimate births, fewer one-person households, and a smaller percentage of older persons in the population. Demographic variables associated with growth from 1951 to 1958 are fewer women in the labour force, higher fertility rates, and smaller percentages of the population of working age.

These 1951-1958 data are interesting in that they vary from standard views of the demographic dimensions of urban growth. Usually one associates lower fertility rates, more women in the labour force, and larger percentages of the population in the working ages with urban growth. The absence of urban size as a variable associated with the basic dimensions of the urban system is interesting because it is assumed to be correlated with a number of other variables as a basic dimension of any urban system.

Mattila and Thompson attempted to investigate some of the causes of urban economic development using a multiple regression model (1968). They regressed measures of economic development of 135 Standard Metropolitan Statistical Areas on a large number of demographic and economic variables selected from census publications. We have listed all of the variables they investigated in Table 33 because of the assumption in regression analysis that all variables affecting the dependent variable that are left out of the model are independent of those included. The inclusion and exclusion of specific variables becomes a crucial determinant of the form the model takes. Table 33 also indicates which variables do not contribute to reduction of unexplained variance in the dependent variables.

The three dependent variables Mattila and Thompson select are measures of the level and distribution of income: median family income (Y_M), the proportion of the population earning less than \$3,000 per year (Y_L), and an index of family income inequality, $\frac{Q_3 - Q_1}{Q_1 + Q_3}$, (Y_I). The standardized regression coefficients derived from these regression equations are (Mattila and Thompson, 1968, p. 65):

$$\begin{aligned} Y_M &= -.47E_L + .30F + .18L_M + .18M_{C/L} + .22M_{60} + .25E_{H4} \\ &\quad + .14P_{60} - .11L_S + .14M_D \qquad R^2 = .87 \\ Y_L &= .55E_L - .31M_{60} - .31F - .19L_M - .11M_{C/L} + .12L_S \\ &\quad - .12E_{H4} + .10NW \qquad R^2 = .88 \\ Y_I &= -.35Y_M + .28NW + .29E_I - .31M_{60} - .08L_F \\ &\qquad R^2 = .89 \end{aligned}$$

Table 33

List of Variables Used by Mattila and Thompson in Multiple Regression Models of Urban Development

Endogenous Variables

Y_M	Median income of families, 1959.
Y_L	Percentage of families with less than $3,000 income, 1959.
Y_I	Income inequality $(Q_3 - Q_1/Q_3 + Q_1)$, 1959.
E_M	Median school years completed, 1960.
E_L	Percentage with six or less years in school, 1960.
E_I	Educational inequality (percentage with six or less years in school plus percentage with four or more years of college, 1960).
Y_G	Percentage change in median income, 1950-60.
E_G	Percentage change in median school years completed, 1950-60.
P_G	Percentage change in population, 1950-60.
L_F	Female, percentage in labour force, 1960.
L_M	Male, percentage in labour force, 1960.

Exogenous (Explanatory) Variables

E_H	Percentage with one or more years of college, 1960.
E_{H50}	Percentage with one or more years of college, 1950.
E_{H4}	Percentage with four or more years of college, 1960.
E_{M50}	Median school years completed, 1950.
F	Percentage foreign-born, 1950.
GS	Growth stability $[(1 + P_G)/(1 + P_{G50/40}) - 1]$.
L_S	Male, self-employed as percentage of total employment, 1960.
L_U	Percentage unemployed, 1960.
$M_{C/L}$	Capital-to-labour ratio (total value added--all employees payroll/total employment, 1960).
M_{4L}	Degree of specialisation within manufacturing (total employment in four largest industries/total manufacturing employment, 1960).
M_{C2}	Capital-to-labour ratio (cumulative plant and equipment expenditures, 1958-60/number of production workers, 1960).
M_D	Percentage employed in durable goods industries, 1960.
M_G	Percentage change in manufacturing employment, 1950-60.
$M_{G50/40}$	Percentage change in manufacturing employment, 1940-50.
M_{60}	Percentage employed in manufacturing, 1960.
MNM	Male net migration rate, 1950-60.
N-S	North (1) or South (0).
NW	Percentage non-white, 1960.
ϕ	Percentage sixty-five years old and over, 1960.
P_{60}	Logarithm of 1960 population.
$P_{G50/40}$	Percentage change in population, 1940-50.
W	All employees payroll/total manufacturing employment, 1960.
Y_{IG}	Intergovernmental inequality (population-weighted standard deviation of median incomes of political subdivisions, 1960).
Y_{I50}	Income inequality $(Q_3 - Q_1/Q_3 + Q_1)$, 1949.
Y_{M50}	Median income of families, 1949.

SOURCE: Mattila and Thompson, 1968, pp. 77-78.

Variation in median family income is explained primarily in terms of proportion of the adult population with less than six years of education (E_L), proportion of the population foreign-born (F), the male labour force participation rate (L_M), the capital-to-labour ratio ($M_{C/L}$), the percentage of total employment engaged in manufacturing (M_{60}), and percentage of the adult population completing four or more years of college (E_{H4}). The two education variables not only measure level of education but presumably also reflect occupational skills and general cultural environment. Mattila and Thompson infer that college graduates not only earn a high income but also contribute to incomes of others through economic leadership, entrepreneurial activities, and general competence in management and technology (1968, p. 64).

The male labour force participation rate (L_M) is the ratio of males in the labour force to the total number of males. It is partly a demographic variable, reflecting the age distribution of males, and partly an economic variable, indicating the availability of jobs for those of working age. The relationship between median family income and male labour force participation is positive, indicating that higher median family incomes prevail in those SMSA's with higher proportions of the age distribution of working age and/or with higher employment rates for those males of working age.

The remaining two factors, percentage of total work force engaged in manufacturing (M_{60}), and the capital-to-labour ratio ($M_{C/L}$), reflect some of the advantages of the manufacturing sector such as the need for more skilled workers, more capital per worker in manufacturing, and higher profits associated with manufacturing industries.

The presence of one of the more powerful predictor variables, percentage of the population foreign born (F), is difficult to assess. Mattila and Thompson suggest that it reflects the migration policy of the United States. International migrants were not allowed into the country unless they were highly skilled. They also suggest that migrants may work harder than native Americans because of work patterns developed in Europe or because of their sense of insecurity in their new country (Mattila and Thompson, 1968, p. 66). A third possible explanation they propose is that contact between the old and new cultures results in a "stimulating and productive interface of cultures" (Mattila and Thompson, 1968, p. 66). They profess discomfort in explaining the presence of this variable but note that it plays a persistent and powerful role in explaining variation in family income.

The second dependent variable, proportion of the population earning less than \$3,000 per year (Y_L), is an alternate measure of level of income. It regresses on the same variables as median family income with approximately the same beta weights, with minor exceptions.

The third dependent variable, a measure of inequality of family income, is associated with median family income (Y_M), non-white as a percentage of total population (NW), educational inequality, which is the sum of the percentages of the adult population with less than six years of school or with one or more years of college education (E_I), and percentage of the labour force engaged in manufacturing industries (M_{60}). Income inequality varies inversely with level of median family income and with percentage of the labour force in manufacturing, while it varies directly with education inequality and with percentage non-white. The relationship of income inequality and education inequality is reasonable given the association of education and income in the previous two regression equations. The association of income inequality and percentage non-white is also acceptable in light of the record of discriminatory procedures of many employers, and the relative poverty of the non-white population in the United States.

A less obvious association is that of level of income and income inequality. Those SMSA's with lower levels of income are also burdened with greater income inequality. Apparently when the distribution of incomes is compared in different urban areas, the greater ranges are seen to occur because of the extension downward of lower incomes. This extension of the population into the lower income levels pulls down the median income. Thus median incomes vary because of differentials at the lower levels of the income scale rather than because of differentials at the upper levels.

The association of percentage of the labour force in manufacturing and income inequality is even less obvious than that of median income and income inequality because the relationship of income inequality and median income has already been explained. After controlling for level of income and level of education, percentage of the labour force in manufacturing continues to account for a significant amount of variation in income inequality. The only possible conclusion is that those SMSA's having large manufacturing sectors tend towards more egalitarian incomes than do SMSA's with small manufacturing sectors.

A study of economic and cultural variation in Canada provides some insight into the demographic correlates of economic development although that was not the intention of the study. D. Michael Ray analyzed 76 socio-economic variables and spatial locators for each of the 229 census counties in Canada using 1961 census data (1969). Although his intent was to map the heartland-hinterland, urban hierarchy, and intermetropolitan axis patterns in Canada, his analysis allows us to observe those demographic factors that are associated with economic growth as well as with the urban hierarchy in Canada. After factor analyzing his data he was able to identify nine socio-economic dimension of Canada, three of which related to economic status. He labelled these three dimensions centre-periphery contrasts, urban-rural contrasts, and metropolitan centres.

The centre-periphery factor focuses on the perpendicular distance of the county from an axis running through New York and Chicago. Three variables which load on this factor make it an economic indicator. The three are an economic disparity index for 1961, another for 1931, and the percentage of the male labour force composed of loggers, fishermen, trappers, and hunters. The economic disparity index was computed from the difference between the retail sales and the expected retail sales based on the population size of the area. The index increased as the ratio of retail sales to population decreased. Toronto had the lowest economic disparity.

The economic disparity index reflects variations in per capita retail sales from one area to another. Reduced per capita retail sales reflect either a limited income of the population with which to make purchases or reduced availability of goods to be purchased. One would assume that it is more likely the reflection of reduced income because when a population has the wherewithal to purchase goods, corporations usually are able to find some means to avail them of desired goods.

The percentage of loggers, fishermen, trappers, and hunters focuses on another aspect of economic development, that of the nature of the economic activity in the region. Regions with relatively large percentages of the labour force in these occupations do not have many secondary or tertiary industries forcing the labour force into less lucrative occupations. These primary activities also require that the region be relatively sparsely settled, denying the region large markets with which to attract industries.

The demographic variables associated with this factor were the male/female labour force ratio and the age structure of the population. As the economic disparity index increased, male domination of the labour force also increased. There were fewer opportunities for women in the labour force because of the economic structure of the region, with fewer manufacturing and service occupations which usually employ women.

Three variables pertain to the age structure of the population: percentage 0-14 years of age, percentage 40-54 years of age, and percentage 55-64 years of age. The two older age groups vary inversely with the economic disparity index while the youngest age group varies directly with it. This indicates a much younger population situated in those regions with greatest economic disparity. High fertility rates and out-migration of persons reaching labour force age would account for this type of distribution. Such an age structure is typical of depressed regions situated within migration range of more advanced regions offering job opportunities to the younger workers.

The urban-rural factor was called an economic indicator because it loaded on a number of economic variables. These included family income, value of housing, rent, and the 1931 and 1961 economic disparity indexes. Urban areas ranked higher than rural areas in all economic variables. The demographic variables associated with this factor included education level and occupational structure of the labour force. Urban areas had a better educated labour force and contained higher proportions of persons in professional, technical, and managerial occupations as well as having higher proportions of craftsmen and production workers. There were higher proportions of females and immigrants in the urban labour force. The age structures differed to the extent that the dependency ratio in urban areas was lower and the percentage of persons 30-39 years of age was higher than in rural areas. All of these contrasts are typical of rural-urban differences and require no further comment.

The third factor was labelled metropolitan centres because it reflected the size, density, and the ethnic composition of the four largest cities in Canada - Toronto, Montreal, Vancouver, and Winnipeg (Ray, 1969, p. 16). It was classified as an economic indicator because of two variables, market potential for 1961 and market potential for 1931. Market potential reflects accessibility to the national market. The highest market potential in 1961 was in Toronto - one-third of all Canadian retail sales were made within one hundred miles of this centre (Ray, 1961, p. 10).

The demographic variables associated with this factor were total population, population density, percentage with Yiddish mother tongue, percentage with Italian mother tongue, and percentage born in Italy. This reflects the concentration of the Jewish and Italian ethnic groups in the largest urban centres in Canada.

The interesting aspect of Ray's study is not the high associations he found between economic and demographic variables, but the lack of association he found between certain economic and demographic variables. The largest number of economic and demographic variables loaded on the urban-rural contrasts factor, indicating that this dimension encompasses most of the economic and demographic variation in Canada. These economic and demographic variables did not load on centre-periphery contrasts nor on the metropolitan centres factor. Thus variation in family income is associated with the rural-urban dichotomy rather than with metropolitan size or distance from the Canadian economic heartland. Variation in education is not associated with metropolitan size nor with the economic heartland; nor is the dependency ratio related to these two dimensions.

Very few demographic or economic variables are related to metropolitan size. The variables that load on this factor are reflections of the size of the large metropolitan centres and of patterns of locational choices of Jewish and Italian ethnic groups.

A demographic characteristic that does not associate with the rural-urban dimension is the age distribution, with the exception of two aspects: the dependency ratio and the percentage of the population aged 30-39 years. Three variables reflecting age structure, percentages aged 0-14 years, 40-54 years, and 55-64 years, are associated with the distance from the heartland. This enhances the heartland-hinterland concept as a useful tool in analyzing the social and economic development of Canada. Not only does it indicate the geographic dimension of economic growth but it also reflects essential demographic processes. Fertility, as measured by percentage of the population aged 0-14, tends to increase as distance from the heartland increases. The association of fertility with rural-urban contrasts and with metropolitan size is overshadowed by its association with distance from the heartland. Fertility is closely tied with norms and values regarding the family institution, which in turn are closely associated with a host of other norms and values relating to modernization.

The other demographic process associated with the heartland-hinterland is migration. This is reflected in differential mortality, which is indicated by the small percentages of the population aged 40-64 found in regions more distant from the heartland. This is the outcome of poorer health facilities in areas further north of the heartland. The age structure can also be interpreted as an indicator of migration patterns. Migration of persons entering the labour force loads primarily on the rural-urban factor as indicated by the predominance of persons 30-39 years of age in the urban population. This is a typical rural-urban phenomenon. However, the shortage of older persons in the more northern Canadian communities suggests one of two possible processes. It may reflect a general migration pattern of older persons, or of adults of all ages from the hinterland toward the heartland. This migration is not discernible in any one age category but only becomes apparent for age groups over 40.

The shortage of older persons in hinterland regions may also reflect a more recent settlement in those areas. These regions have only begun to be settled in more recent times and the ages of the population moving into these regions have been typically younger. The age structure thus reflects a younger, more vigorous population. The population of the Northwest Territories has increased rapidly in recent years as the result of expansion of mining industries into this region and a concurrent growth of government services. Similar expansion has occurred in the northern parts of the three prairie provinces. The movement of younger families into the hinterland also contributes to the higher proportions of children aged 0-14 found there. Recent migration of families into the hinterland combined with the higher mortality rates of the resident populations are probably the prime factors in variations of the age structure of the heartland-hinterland populations.

In summary, Ray's study provides some interesting insights into spatial patterning of economic and demographic characteristics. Economic variation is best explained in terms of urban-rural contrasts. Variation in the age structure of the Canadian population is best explained in terms of the heartland-hinterland dichotomy. The age structure in turn indicates patterns of fertility, mortality, and migration. Migration rates for age groups 20-39 appear to occur along rural-to-urban patterns, but a pervasive migration of persons of all ages occurs from the heartland to the hinterland as the result of government policies and mining activities. None of these demographic processes appears to be associated with urban size.

We may benefit by returning to the study conducted by Leslie King (1966) and discussed in pages 217-233 in terms of his city groupings and our measure of urban economic development. Now we shall look at the demographic dimensions which he investigated. King analyzed Canadian urban dimensions in 1951 and 1961. He extracted 12 components from the 52 variables from the 1951 data, including several demographic measures. Most of the demographic variables loaded on the first two components.

His first component is an index of the youthfulness of the female population in non-manufacturing and less densely populated cities. Cities of British Columbia, the Prairies and the Maritimes ranked highest on this component while cities in Quebec ranked lowest (King, 1966, p. 209). Regionalism is apparently a factor in this dimension. The number of immigrants in the preceding decade also had a high loading on this component, which may explain the high proportion of younger women, since international migration is age-selective. King unfortunately did not include variables in his input relating to the age structure of the male population.

Most of the economic variables loaded highest on this component indicating perhaps the influence of low wages paid to immigrants. The large percentage of young women can be attributed to internal migration as well as to international

migration, although King provides no data to verify this suggestion. If this is the case, then the high rates of female in-migration can be attributed primarily to job opportunities for females. Since females are paid lower wages than males, the preponderance of females would also account for the poorer economic indicators of these cities.

The second component has a socio-economic character.

Variables measuring educational attainment have high positive loadings while those indicating the French-Canadian influence have high negative loadings. Cities ranking high on this component are located in British Columbia and the Prairies while cities ranking low on it are found in Quebec. Two age structure variables have high loadings on this component, reflecting a low proportion of the population in the younger age brackets and a high proportion over 65 years of age. High loadings of per capita retail and wholesale sales indicate fairly viable economies in these cities, even though measures of isolation show that they are not proximate to other cities.

These two components provide us with some generalizations for cities in Canada in 1951. The first component taps cities outside of Quebec and Ontario with lower levels of economic development. The most surprising demographic aspect of these cities is the high proportion of females aged 15-39. This reflects a high rate of in-migration, which is usually associated with cities with more buoyant economies. The male age distribution appears to parallel that of the female in that the percentage of the total population in the active labour force also has a high loading on this component. These cities have a favourable dependency ratio, yet do not benefit economically. Their labour forces are characterized by relatively fewer persons involved in primary industries, and in industries of transportation, retail, wholesale, finance and insurance. The deficiency of financial and insurance activities indicates that the cities are small, because such industries are typical of metropolitan centres.

The economic condition of the cities may be a reflection of regionalism and of their size. They are smaller western cities whereas the most economically advanced are Ontario cities and larger cities outside Ontario. Their age structure reflects high rates of rural-urban migration. Western Canada began to industrialize during the Second World War, encouraging a considerable amount of migration to cities from rural areas. Thus, the cities have favourable age-sex profiles without the economic advantages.

The second component is also regionally affected in that it taps cities in British Columbia and the Prairies. These cities are isolated and have low manufacturing profiles but nevertheless enjoy a fairly high economic standing. The populations are well-educated and reflect this in their demographic age structure; fertility is low while the proportion of older persons is high. The unusual demographic correlate is the dependency ratio. It is lower in the low-income cities associated with the first component than with the higher-income cities associated with the second component. These cities have the lowest proportion aged 0-14, indicating that their adverse dependency ratios are inflated by large proportions in the older age groups. These cities are not as small as the cities of the first component, reflecting the tendency for retiring persons to move to larger cities having more and better services.

King replicated his study using 1961 data. He found a number of structural changes in the Canadian urban system over the ten year period. One apparent change was a separation of demographic and economic dimensions. The first component from the analysis of 1961 data was primarily economic. It contained no high loadings from demographic variables, which loaded on the second, third, and fourth components. The second component was clearly identified with the Quebec population structure (King, 1966, p. 218). A large percentage of the female population was between the ages of 15 and 40, a small proportion of the total population was over 65 years of age, while a large proportion was under 15.

The third component had a low dependency ratio, primarily because of low percentages of the population under 15. The fourth component had high loadings on the feminity ratio. The smaller cities in Ontario ranked high on this dimension.

King's study points out the hazards in making generalizations about demographic and economic correlates in the Canadian urban system. He found some association between large proportions of females aged 15-39 and low economic status for smaller cities outside of Ontario and Quebec. At the same time the age structure in western cities was found to be associated with economic status. Cities ranking higher in economic status had fewer young persons and larger percentages of old persons in their populations.

The weak association between economic status and demographic characteristics is further weakened in 1961. Demographic variables load on dimensions that are independent of economic variables.

What one can surmise from this and Ray's studies is that a number of intervening variables must be considered in analyzing demographic and economic interrelationships. The functions of cities, their regional location and their positioning according to the heartland-hinterland concept all play key roles in determining the economic and demographic conditions of the cities.

Two regional studies verify the tenuous association between demographic and economic characteristics in Canada. Hodge analyzed economic, demographic, and sociological variables pertaining to all incorporated areas in Saskatchewan with the exception of the four largest cities (1965). The only demographic correlate of economic potential was urban size. The age structure was independent of this dimension. Population density and population growth were correlated indicating that rapid growth in incorporated areas tends to occur within the corporate boundaries, causing an accompanying increase in density. But these two variables were not associated with age structure, nor with population size and economic status.

A second regional study evaluated sociological, demographic and economic variables in the Alberta urban system (Stafford, 1975). Once again, the only demographic variable associated with economic standing was population size. Age structure did not correlate with economic variables with one exception. Cities with larger proportions of old people had fairly distinctive industrial profiles. Large percentages of the labour force were engaged in trade, finance, and service industries and in managerial and professional occupations. This makes Alberta somewhat unique in that the traditional metropolitan functions of finance, insurance, and real estate are not centred in the largest urban areas. Instead these functions are disproportionately located in some of the medium-sized towns in central Alberta having specialized functions such as agricultural stations, and corrective and health institutions. The large number of older persons in these towns suggests that they serve as retirement areas for many of the rural people in central Alberta.

A Comparison of the Three Theories

The three theories of regional development, growth pole theory, export base theory, and sector theory, were conceived and nurtured by economists and geographers. All three are consequently deficient in an important aspect: they omit the demographic dimension of economic growth. To state that a region is experiencing growth or development is to state that the population within the boundaries of that region is undergoing change. The population may be increasing or decreasing, or the attributes of that population may be changing, but only under the most unusual circumstances would a region develop without there being concomitant change in the population. Such a situation would be an influx of capital into an area with no alterations in the labour force. However this would be short-lived because the capital would create new jobs and require specific skills of the labour force. The creation of new jobs would increase personal income within the region, creating a new distribution of wages which itself is a demographic change. Any increment in capital or labour would result in variation in the demographic and ecological make-up of the region in a very short time.

The question we wish to address in this section is one regarding the demographic and ecological viability of each of the three theories of regional development. Not one of them explicitly recognizes these dimensions but the demographic and ecological implications are perhaps more viable in one than in the others. The procedure will be to trace the demographic implications of each theory in turn, to contrast the differences in the three theories, and to evaluate the viability of the three theories in terms of the ecological complex.

Export base theory traces the development of a region from its original subsistance stage. Gradually the region advances economically on the basis of the growth of exports of staple commodities to more advanced regions. The rise of export industries attract a number of locally based industries to serve the local market. Eventually these are supposed to develop to the extent that they also export their services and commodities to other regions.

The key to the growth of the region is investment, and in order for the region to develop investment must occur during the early stages of development when the economy is still largely at the subsistence level. This may be untenable because we know that investors are more comfortable making their investments in areas with large local markets and diversified economies, providing the investors with greater confidence that they will realize returns on their investments. However, the condition of investment in a subsistence region becomes acceptable when we realize export base theory is not supposed to be applicable to any subsistence region but only to those containing an exportable staple commodity for which there is an external demand. With adequate demand, there is certain to be investment regardless of the level of economic activity already in the region.

At this point, export base theory could perhaps be criticised for being politically naive. In the present era of demand for native rights in northern Canada and concern with imperialist exploitation in underdeveloped countries, a region or country with subsistence economies and exportable staples is always suspicious of foreign investments. The political forces opposing such investment may play a crucial role in the rate of investment and the effectiveness of that investment in contributing to regional development. However, when export base theory was evolved during the 1930's and 1940's such political concerns were not as prevalent as they are today.

Ideological aspects aside, export base theory can also be criticised for ignoring the crucial sociological and demographic processes that would necessarily follow investment in a region with a subsistence economy. The addition of capital stock requires certain attributes, skills, and knowledge which the encumbent population is almost certain to lack. Attainment of these attributes would require either selective in-migration of persons who are by definition different from the resident population, or the selective training of certain members of the region so that they are able to manage and maintain the influx of technology that necessarily accompanies additional capital.

The alternatives of educating local persons or attracting migrants are not mutually exclusive; the education approach would require in-migration of teachers and workers with the appropriate skills to construct the educational institutions. Some in-migration is necessary and usually the migrants will have higher incomes than most of the local residents resulting in their having a considerable impact on the economic and social institutions in that area. They will demand better standards of housing, leisure and services than prevail in subsistence communities. The new social and economic order that accompanies high levels of in-migration into a region results in a certain amount of alienation on the part of the residents. They will either withdraw from activities which portend economic growth or endure the process of acculturation which allows them to adapt to the new conditions. In either case there is a disruption in the old social order and a certain amount of stress which has repercussions on the ensuing technological developments. These repercussions are too crucial to be ignored by the theory. They have had profound effects on the development of the Canadian north, where drunkenness, broken homes, illegitimacy, and crime become major social problems. These social problems are obvious deterrants to a rapid economic development of the North. No man can learn new trades in technical school or function adequately in an airport control tower when he is subject to these forms of maladjustment.

An important paramater in the sociological context of the first stage of growth in export base theory is time. How long does it take for the region to move from a subsistence economy to one based on staple exports? The shorter the transition period the more crucial is the social dimension.

A very slow transition would allow the residents ample time to adjust to the new conditions and to assimilate the ideas imported by the in-migrants.

No mention is made of the time span involved in this process. A clue to the time scale can be gleaned from the applications of export base theory. It is supposed to be characteristic of the process of development of the United States and Canada and of regions therein. In this case the time scale covers several decades, but more crucial, the development takes place in a region which is virtually unpopulated, or in which the indigenous population plays no role in the development of an export base. If these unpopulated regions are the ones to which export base theory pertains, then conditions under which it is applicable are very narrow, and the theory is out of date. This condition is never specified in the literature, so one would assume that the theory can also be applied to regions of sparse population, and in fact that condition has been specified.

Assuming an exportable staple in an under-populated region and assuming that the influx of capital and labour does not create serious political or social problems, the next necessary condition for growth is that the resource endowment have location-associated forward linkage effects. That is, industries which use the resource as inputs will be attracted to the region to reduce transportation costs. The factor that makes this step probabilistic is a demographic one. Most industries with forward linkages from primary industries are manufacturing industries. But in most cases these industries will select regions with large populations for their sites. They do this because they require skilled labour and populous areas offer a greater choice of labour skills than a sparsely populated area. They also select such sites because transportation costs dictate that they must locate in or near a large market. In most cases the raw materials can be shipped from the source to the manufacturing site much more cheaply than the finished product can be shipped from the manufacturing site to the market. Raw materials do not need to be carefully crated nor do they take up as much space as the finished product. The manufacturing firm is dependent on a number of other linked industries for inputs to its production process and the manager prefers to locate where these are situated, which again is usually in a more populous region. If he chooses to locate the firm near the source of raw materials he must import all other inputs and then transport the product a considerable distance. If competitors are located in a market area with a ready supply of labour and materials, they are able to sell at lower prices and maintain higher profits. They will also have access to a larger, more diversified labour pool.

The necessary condition for attraction of manufacturing enterprises is a large local population. This may require a lengthy time period from the date when the region became an exporter of raw materials because there is no apparent reason for a large population to migrate into the region. The region would have to rely on natural increase for its source of population growth. Canada exemplifies this condition. It has been an exporter of staples for three hundred years but it is still a sparsely populated country. The only extensive manufacturing that has developed has been in southern Ontario where access to the large markets in the eastern United States and Europe made these industries viable.

The prime mover in export base theory is demand, which is a function of the external market for the resource endowment of the region. The market depends upon the distribution and size of the population outside the region and its location relative to the region. Thus the population, its size, composition and patterns of consumption determine the economic viability of the developing region. Export base theory should be criticized for glossing over the attributes of the population which determine the viability of the theory itself.

The theory can also be criticised for omitting any analysis of the change in the labour force, both quantitative and qualitative, within the region which contributes to the development of the region and which is a necessary condition for attracting manufacturing firms to the area. The theory is remiss for its omissions rather than for distortion of the demographic factors that contribute to growth.

Sector theory describes regional development as a function of differential elasticity of demand for primary, secondary, and tertiary commodities and services. This results in shifts in the major production of a region from primary to secondary and from secondary to tertiary industries. An essential concomitant factor of growth is reduced transfer costs resulting from an improved transportation system. Sector theory specifies that originally the region has a subsistence agricultural economy. The economy evolves as some specialisation in agricultural production occurs as the result of trade in primary products with other regions. This is made possible through improvements in transportation.

Sector theory begs the question why persons engaged in subsistance agriculture should convert their economy to an industrialised one. If the theory is viable, it should provide a prescription for the economic development of most of the underdeveloped countries in the world today. It specifies that agricultural specialisation and industrialisation will occur before overpopulation perpetuates poverty and cancels any effects that investment might have on the region's growth. Rather, it specifies that the region begins as a sparsely populated locality and then describes development in terms of increasing population, but not overpopulation. The increasing agricultural population is supposed to reduce per capita returns, causing agricultural investment to be diverted to industrial purposes and excess agricultural workers to take up manufacturing activities. The possibility of overpopulation contributing to economic stagnation is ignored even though this process appears to be the prevalent one in the third world.

Many of the criticisms of export base theory can be repeated regarding sector theory. They both recognise a need for industrialisation while ignoring the social and demographic factors that must accompany the upheavel in the economic structure of the labour force. Are the new skills imported or trained? What effects do these have on the traditional institutions and ways of life that have prevailed among the indigenous population? Why are manufacturing firms willing to forsake other locational advantages to move into a relatively underdeveloped region?

Both sector theory and export base theory are guilty of omitting social and demographic dimensions of industrialisation but sector theory exceeds these sins of omission. It describes the development of a subsistence agricultural region into a specialised agricultural and industrial economy as a function of both differential elasticity of demand and population pressure. Empirical observations suggest that an increasing population has the opposing effect of differential elasticity of demand in that it absorbs savings that would otherwise be converted into investment. This places sector theory in the position of being falsified by empirical evidence.

Growth pole theory diverges from the others in that it does not begin with a sparsely populated, subsistence economy. No statement is made about the initial structure of the economy. The only requirement is that situated within that region there be a rapidly growing, technologically advanced industry. Its growth stimulates growth in other industries through backward and forward linkages and provides a local environment that encourages innovation. The growth pole is supposed to be located in a growth centre which is a point of agglomeration containing high level tertiary industries, a large labour market, and strong linkages with the national economy.

A number of demographic dimensions are implied here although none are specified. The fact that the growth centre has agglomeration advantages indicates that it is a fairly large city although not so large that further growth would lead to diseconomies of scale. There is little agreement on the size of the growth centre although clearly if a theory is to be developed it must specify the range within which the population of a growth centre must be. These ranges would vary according to the location of the growth centre relative to the distribution of the population within and outside the country. The required size of a growth centre would also be a function of the economic development of the national economy. The minimum size of a growth centre in the United States would necessarily be larger than one in Tanzania. A number of conditions that must be specified in defining a growth centre would include the relationship between size and rate of growth for a given state of development in a given sociocultural system.

Growth pole theory is fairly explicit in describing the nature of the population growth. The growth centre will attract industries requiring more persons and higher-level skills than are available in the centre. This requires a considerable amount of in-migration and the displacement of workers in slow-growth industries. The question of where they come from is never discussed although some mention is made of attracting young, energetic workers from the hinterland at the expense of further deterioration of that area.

A significant proportion of the migrants must be highly skilled and well educated because they must initiate innovations which contribute to the technological primacy of the growth centre. Such professional and managerial personnel must come from other major urban centres and must be well trained and experienced in their fields. This implies that the growth centre is one of a system of cities in a well developed economy. The country must have the means to educate and train a significant portion of its labour force for entrepreneurial functions.

In summary, growth pole theory is remiss in specifying the demographic characteristics of the growth centre, although these characteristics are crucial to the theory. The theory assumes a ready supply of technological and entrepreneurial expertise outside the growth centre. The attraction of these personnel then depends upon the activity of a growth pole located in the growth centre. The growth centre must be large and technologically sophisticated so that there are not the problems of social adjustment that are implicit in the descriptions of development offered by the other two theories.

The ecological complex presents an alternative method of comparing the three theories. A theory of regional development is an ecological theory. The difficulty with present theories of regional development is that they select a limited number of variables in their framework, mainly economic variables, and use a deterministic approach wherein growth is promised upon the correct manipulation of specified economic variables. Human ecology provides a much more comprehensive view of the real situation and may in the end contribute the synthesis necessary for a theory of comprehensive regional development, something that is lacking in existing theories.

An ecological approach would involve the examination of the four clusters of variables - population, organisation, environment and technology. The researcher would observe how they interact in times of growth and stagnation and thus he would identify the causes and consequences of growth. The most difficult task would be that of determining the necessary and sufficient conditions, as distinct from concomitant conditions, of growth.

The terms of export base theory can easily be expressed in the form of the ecological complex. The environmental qualities of the region include a sparse population, a valuable resource, and a set of regions which are more fully developed, and within which there is demand for the endowment of the undeveloped region. Note that the ecological complex provides us with an analytical framework. The four components, population, environment, social organisation, and technology are not separate and distinct entities but for purposes of analysis they can be so separated. Thus when we discuss the environment, we must keep in mind that it is inseparable from the other components.

Population, technology, and social organisation also play their parts in the initial stages of development, but environment is most essential. The region must have an exportable resource before export base theory can be applied. The resource must be extractable and other regions must have the capital, the techniques and the desire to extract it.

Given the environmental prerequisites, the other three components of the ecological complex come into focus. The region must be underpopulated. No mention is made in the literature about the social characteristics of the population but it is implied that educational levels and living standards are fairly low; otherwise there would not be a need for development. The most important aspect of the population component is that the population be small. In such a case, variations in the sociological profile of the population will have a much smaller impact than when the population is relatively large.

The technology and social organisation of the initial conditions of the region are also neglected in the literature. Once again, we can assume that these are relatively primitive given the need for development of the region. The economy in the region is at a subsistence level, implying a very simple technology. A sparse population implies a simple social organisation. These details are not described in export base theory but would be if the theory were couched in terms of the ecological complex.

Although the attributes of the environment are crucial in the initial stages of development, the interaction of the other three components determines the economic growth of the region. Technology is imported in the form of extractive procedures and transportation facilities to allow the valued resource to be exported. However, it cannot be put into operation unless the labour force is made available to apply and administer the new techniques. The technology and social organisation play equally important roles in moving the region from a subsistence to an exporting economy.

The role of the environment diminishes as technology and social organisation raise the living standards and profits of the region. However these in turn must be partially replaced by population as the prime mover if regional growth is to be sustained. Few resources are renewable, and those that are, such as trees, do not have the interregional demand to provide the vehicle for continued growth that other resources, such as minerals, have. Thus the export base must eventually shift to manufactured products and services. At this stage, the key component is population. Other ecological components must perform at a level that attracts manufacturing industries to the region but these will locate in the region only if an adequate population is there. The population will grow only if jobs and amenities are made available by the existing technology and social organisation. The role of the environment takes a notable shift in that now it becomes a provider of amenities to attract migrants.

The interaction of all four ecological components, but especially social organisation, technology, and population combine to sustain growth during the middle stages of development in the region. But eventually the crucial component becomes population. A large and skilled population provides innovations, and technological expertise to allow the region to become an exporter of these services. This represents the final stage of development in export base theory.

When the ecological complex is applied to sector theory a number of similarities and distinctions between this theory and export base theory become apparent. Proponents of sector theory emphasize the role of improved transportation as being the essential feature of development. This places technology at the front of the other ecological components in sector theory. The region relies on a subsistence agricultural economy initially, so the setting invokes all four components, a sparse population, a simple social organisation, an undeveloped technology and an agrarian environment. However, this is only the setting. The key requisite in sector theory is an improved transportation technology while the key requisite in export base theory is the existence of a valuable natural resource.

An equally crucial component, but one that plays a role after technology has been in effect is population. This factor combines with improved elasticity of demand in the manufacturing sector to shift the labour force into manufacturing. Thus we have a direct causal relation between population and social organisation. The increase in population leads to a more complex social organisation.

During the middle stages of growth the dominant manufacturing process shifts from consumer goods to capital goods. Capital-goods industries are geared to large-scale production, allowing them to benefit from economies of scale. This involves a considerable number of adjustments in the production process. New materials are required and more elaborate organisational techniques are invoked involving greater specialisation, more linkages, and a higher input of capital. The shift in technology is crucial to further development of the region. Thus technology, which has been labelled the basic component throughout the developmental process, rises to the forefront again after briefly giving way to population as a driving force in an earlier stage of development.

The role of environment is never specified in sector theory and presumably plays a minor role. Within the framework of the ecological complex, this factor would at least be spelled out whereas it is never mentioned in the regional theory of growth. Variation in technology, and to a lesser degree, population, is the major vehicle of growth. The only aspect of population that is explicitly considered is size, although as industrialisation takes place, a number of changes in the characteristics of the population are implied. The labour force becomes technologically sophisticated as it must in any industrialising region. In conjunction with these changes, the social organisation of the region must necessarily adapt to the technological advancement, but sector theory gives this component only minimal attention.

In terms of the ecological complex, sector theory differs considerably from export base theory. Sector theory ignores environment, and focuses on the interaction of technology, population, and social organisation. Technology is the prime mover, with population contributing to the shift from an agricultural to a simple manufacturing economy, after which technology dictates the process of change to more elaborate forms of manufacture. Export base theory begins with a specific environmental attribute. Further economic maturation is a function of the interaction of technology and social organisation, and eventually population.

Growth pole theory could be described in terms of the ecological complex, and indeed, the refinement of the theory may hinge on its taking on a broader framework such as the ecological complex provides. Growth pole theory begins with technology and so this set of variables would correspond to the centralized allocative mechanism around which all other variables revolve. Introduction of new technologies would cause a shakeup in the encumbent technological structure; firms would be forced to adapt to the new competition or succumb. The new technological structure creates a labour vacuum, attracting certain segments of the population and repulsing others. Persons with skills suited to the new technology move into the area while those whose skills are not suited are forced to move elsewhere or upgrade their skills. The standards of skills must improve in a growth centre - this is a necessary condition. Changing from one type of skill to an equivalent but different one does not encourage growth. It may for a short period as the market shifts but eventually growth can occur only when lower-order skills are replaced by those of a higher order. The inflow of higher skills must meet the demand and at times must exceed it if wages are to stay competitive and in order that a reserve pool of skilled labour be available to meet fluctuating demands and act as a pole of attraction to other firms. Thus population variables enter into the complex, generated by technological change and in turn contributing to further technological change. It is only through the process of mutual attraction between skilled labour and appropriate industries that agglomeration occurs - another necessary condition of a growth centre.

So far we have discussed two of the four sets of variables that make up the ecological complex. However the restructuring of technology and population combined with their increased scale has a profound effect on the other two variables. With respect to organisation an increased population requires changes in the spatial organisation with greater territorial differentiation. Distinctive residential areas evolve as the population increases and its socio-economic structure diversifies. Demarcation of residential, business and recreation areas becomes more definite.

The morphology of job opportunities affects the age-sex structure of the population - more women enter the labour force, fewer children are born, and younger persons with increased levels of job training restructure the population. A proliferation of occupations occurs making the growth centre more attractive to firms dependent on skilled labour and consequently more attractive to potential in-migration. Thus the three variables, population, technology and organization are in a continuous, dynamic state of interaction, each one contributing to the growth of the other two.

In conjunction with change in organisation is the change in environment in that the organisation is part of the environment. Beyond the organisation is the growing number of contacts with other cities, with extra-regional corporations, and with regional, national and extra-national governments. Not only are the number of contacts increased but their quality and diversity are also elevated by the improved system of transportation and communication which accompanies economic and demographic growth. Thus technology and population as well as organisation contribute to the quality and dimension of the environment.

The resulting environment acts as a catalyst to innovation which is an important component of growth pole theory. Innovations contribute to a richer technology and a more flexible and effective organization all of which combine to attract more persons to the area.

In summary, human ecology can broaden the perspective inherent in growth pole theory. The latter has failed to predict because it has focused primarily on technology as the propulsive force in growth. It has recognized population - and to a lesser extent, organisation and environment - as a contributing factor in growth. But it has placed the latter three in a position dependent upon technology, rather than recognizing that these three interact with each other and with technology to play a propulsive role in the ultimate growth of a region.

Summary of Dissertation

The major concern is the identification of factors leading to differential urban growth within a system of cities. Urban growth is a recent phenomenon that has permeated every dimension of human society. Causes of urban growth are primarily economic. Technological and commercial development, fluidity of the social class structure, population increase, and industrialization are considered important factors in the proliferation of cities in the past 100 years.

No adequate theory of differential urban growth exists. Export base theory, central place theory, and communications theory address this problem but fall short of an adequate explanation. Alternate sources of explanation may lie in regional science.

Three theories of regional growth are applied to data from the Canadian urban system to evaluate their contributions to differential urban growth. The three are sector theory, export base theory, and growth pole theory. Sector theory states that a region develops as the major components of its economy shift from primary to secondary and from secondary to tertiary industries. According to export base theory, a region develops because it contains a commodity which is in demand outside the region. Growth pole theory attributes development to agglomeration, industrial linkages, and entrepreneurial activity.

Variables that are considered to be essential to each theory are operationalized. The corresponding data for the forty-six largest cities in Canada are regressed on migration and wage data using a stepwise multiple regression model. The results indicate that the variables representing export base theory contribute very little to reduction of the unexplained variance in the dependent variables. Sector theory contributes more, and growth pole theory most, to reduction of this variance. Two variables from growth pole theory, agglomeration and profits, are most powerful in accounting for variation in wages and migration. The most powerful variable from sector theory is percentage of the manufacturing labour force engaged in capital-goods industries.

A theory of differential urban development in Canada must extend beyond traditional economic variables. Other factors, as indicated by studies of Canada's urban system include regional characteristics. Associated with regionalism is the heartland-hinterland concept and the economic shadow concept, both of which favour cities in southern Ontario. The financial central place functions of some large cities, particularly Toronto and Montreal, must also be considered. Finally the advantages of Metropolitan Areas and capital cities play a role in differential urban development.

Such a theory should also take into account the demographic dimensions of social and economic change. Urban development requires the upgrading of the skills of the urban population. This is brought about by means of in-migration and education. The sources and talents of the migrants as well as the difficulties in adjustment of the resident population must be inculcated in a theory of urban development.

REFERENCES

Anderson, J. 1968. "Economic Base Measurement and Changes in the Base of Metropolitan Edmonton, 1951-1961", The Albertan Geographer 4:4-9.

Andrews, Richard B. 1954a. "Mechanics of the Urban Economic Base: The Problem of Base Measurement", Land Economics 30:52-60.

Andrews, Richard B. 1954b. "Mechanics of the Urban Economic Base: Special Problems of Base Identification", Land Economics 30:260-269.

Andrews, Richard B. 1954c. "Mechanics of the Urban Economic Base: The Problem of Base Area Delimitation", Land Economics 30:309-319.

Andrews, Richard B. 1955a. "Mechanics of the Urban Economic Base: Causes and Effects of Change in the Base Ratios and the Ratio Elements (I)", Land Economics 31:144-155.

Andrews, Richard B. 1955b. "Mechanics of the Urban Economic Base: Causes and Effects of Change in the Base Ratios and the Base Ratio Elements (III)", Land Economics 31:361-371.

Barratt, T. R. 1971. "An Analysis of Fertility Trends in Ontario", Ontario Economic Review May-June, Ontario Economic Council.

Bauer, P. T., and E. S. Yamev. 1951. "Economic Progress and Occupational Distribution", Economic Journal 61:741-755.

Baxter, Clive. 1966. "Rich and Poor Regions: The Gap Isn't Narrowing", Financial Post, January 8.

Bereiter, Carl. 1963. "Some Persistent Dilemmas in the Measurement of Change" in Chester W. Harris (ed.), Problems in Measuring Change. Madison: The University of Wisconsin Press.

Berry, Brian J. L. 1964a. "Cities as Systems within Systems of Cities" in John Friedmann and William Alonso (eds.), Regional Development and Planning. Cambridge: The M.I.T. Press.

Berry, Brian J. L. 1964b. "City Size Distributions and Economic Development" in John Friedmann and William Alonso (eds.), Regional Development and Planning Cambridge: The M.I.T. Press.

Berry, Brian J. L. 1971. "Internal Structure of the City" in L. S. Bourne (ed.), Internal Structure of the City. Toronto: Oxford University Press.

Berry, Brian J. L., and Allen Pred. 1965. Central Place Studies: A Bibliography of Theory and Applications. Philadelphia: Regional Research Institute.

Berry, Brian J. L., and Frank E. Horton. 1970. Geographic Perspectives on Urban Systems. Englewood Cliffs, N.J.: Prentice-Hall, Inc.

Blalock, Hubert M. Jr. 1968. "Theory Building and Causal Inferences" in Hubert M. Blalock Jr. and Ann B. Blalock (eds.), Methodology in Social Research. New York: McGraw-Hill Book Company.

Blumenfield, Hans. 1955. "The Economic Base of the Metropolis", Journal of the American Institute of Planners 21:114-132.

Bohrnstedt, George W. 1968. "Observations on the Measurement of Change" in Edgar F. Borgatta (ed.), Sociological Methodology 1969. San Francisco: Jossey-Bass Inc.

Borts, George H. 1960. "The Equalization of Returns and Regional Economic Growth", The American Economic Review 50:319-347.

Boudeville, J-R. 1966. Problems of Regional Economic Planning. Edinburgh: The University Press.

Bourne, L. S. 1974. "Urban Systems in Australia and Canada: Comparative Notes and Research Questions", Australian Geographical Studies 12:152-172.

Bourne, L. S., and R. D. MacKinnon (eds.). 1972. Urban System Development in Central Canada. Toronto: University of Toronto Press.

Burgess, Ernest W. 1929. "Urban Areas" in T. V. Smith and L. D. White (eds.), Chicago: An Experiment in Social Science Research. Chicago: University of Chicago Press.

Casetti, E., and R. K. Semple. 1968. "A Method for the Stepwise Separation of Spatial Trends", Michigan Inter-University Community of Mathematical Geographers, Discussion Paper, number 11, Department of Geography, University of Michigan, Ann Arbor.

Casetti, E., L. J. King, and J. Odland. 1971. "The Formalization and Testing of Concepts of Growth Poles in a Spatial Context", Environment and Planning 3:377-382.

Coale, Ansley J. 1960. "Population Change and Demand, Prices, and the Level of Employment" in National Bureau of Economic Research. Demographic and Economic Change in Developed Countries. Princeton: Princeton University Press.

Coleman, James S. 1968. "The Mathematical Study of Change" in Hubert M. Blalock Jr. and Ann B. Blalock (eds.), Methodology in Social Research. New York: McGraw-Hill Book Company.

Cox, Harvey. 1965. The Secular City. New York: The Macmillan Company.

Crowley, Ronald W. 1973. "Reflections and Further Evidence on Population Size and Industrial Diversification", Urban Studies 10:91-94.

Darlington, Richard B. 1970. "Multiple Regression in Psychological Research and Practice" in Emil F. Heerman and Larry A. Braskamp (eds.), Readings in Statistics for the Behavioral Sciences. Englewood Cliffs, New Jersey: Prentice-Hall, Inc.

Darwent, D. F. 1969. "Growth Poles and Growth Centers in Regional Planning: A Review", Environment and Planning 1:5-32.

Davis, Kingsley. 1965. "The Urbanization of the Human Population", Scientific American 213:41-53.

Davis, Kingsley. 1972. World Urbanization 1950-1970 Volume II: Analysis of Trends, Relationships, and Development. Population Monograph Series No. 9. Institute of International Studies, University of California, Berkeley.

Denton, Frank T., and Sylvia Ostry. 1967. Historical Estimates of the Canadian Labour Force. 1961 Census Monograph. Catalogue Number 99-549. Dominion Bureau of Statistics. Ottawa: The Queen's Printer.

Denton, Frank T., and Byron G. Spencer. 1970. "Macroeconomic Effects of Population Change: Some Exploratory Simulations", Working Paper No. 70-12. Department of Economics, McMaster University, Hamilton.

Dillard, Dudley. 1948. The Economics of John Maynard Keynes. New York: Prentice-Hall, Inc.

Dominion Bureau of Statistics. 1953. 1951 Census of Canada. Volume IV. Ottawa: The Queen's Printer.

Dominion Bureau of Statistics. 1954a. 1951 Census of Canada. Volume VII. Ottawa: The Queen's Printer.

Dominion Bureau of Statistics. 1954b. 1951 Census of Canada. Volume VIII. Ottawa: The Queen's Printer.

Dominion Bureau of Statistics. 1954c. General Review of the Manufacturing Industries of Canada, 1951. Catalogue Number 31-201. Ottawa: The Queen's Printer.

Dominion Bureau of Statistics. 1960. Canadian Life Tables. Catalogue Number 84-510. Ottawa: The Queen's Printer.

Dominion Bureau of Statistics. 1963a. 1961 Census of Canada. Catalogue Number 94-504. Ottawa: The Queen's Printer.

Dominion Bureau of Statistics. 1963b. 1961 Census of Canada. Catalogue Number 94-518. Ottawa: The Queen's Printer.

Dominion Bureau of Statistics. 1963c. 1961 Census of Canada. Catalogue Number 94-519. Ottawa: The Queen's Printer.

Dominion Bureau of Statistics. 1963d. 1961 Census of Canada. Catalogue Number 94-520. Ottawa: The Queen's Printer.

Dominion Bureau of Statistics. 1964a. 1961 Census of Canada. Catalogue Number 94-521. Ottawa: The Queen's Printer.

Dominion Bureau of Statistics. 1964b. 1961 Census of Canada. Catalogue Number 97-502. Ottawa: The Queen's Printer.

Dominion Bureau of Statistics. 1965a. General Review of the Manufacturing Industries of Canada, 1961. Catalogue Number 31-201. Ottawa: The Queen's Printer.

Dominion Bureau of Statistics. 1965b. 1961 Census of Canada. Catalogue Number 97-518. Ottawa: The Queen's Printer.

Dominion Bureau of Statistics. 1966a. 1961 Census of Canada. Catalogue Number 94-532. Ottawa: The Queen's Printer.

Dominion Bureau of Statistics. 1966b. 1961 Census of Canada. Catalogue Number 94-551. Ottawa: The Queen's Printer.

Dominion Bureau of Statistics. 1966c. 1961 Census of Canada. Catalogue Number 97-512. Ottawa: The Queen's Printer.

Dominion Bureau of Statistics. 1968. Manufacturing Industries of Canada, Section E, Prairie Provinces, 1964. Ottawa: The Queen's Printer.

Draper, N. R., and H. Smith. 1966. Applied Regression Analysis. New York: John Wiley & Sons, Inc.

Duncan, O. D., and A. J. Reiss. 1956. Social Characteristics of Urban and Rural Communities, 1950. New York: Wiley.

Duncan, O. D., W. R. Scott, S. Lieberson, B. Duncan, and H. H. Winsborough. 1960. Metropolis and Region. Baltimore: Johns Hopkins Press.

Duncan, Otis Dudley, Ray P. Cuzzort, and Beverly Duncan. 1961. *Statistical Geography*. Glencoe, Illinois: The Free Press.

Ebel, Bernd. 1973. "Interindustry Linkages and the Agglomeration of Manufacturing Industries in the Canadian Economy", unpublished Ph.D. dissertation, Department of Economics, University of Alberta, Edmonton.

Eddington, Arthur S. 1933. *The Nature of the Physical World*. New York: The Macmillan Company.

Financial Post, January 15, 1966.

Fisher, Allan G. B. 1933. "Capital and the Growth of Knowledge", *Economic Journal* 43:379-389.

Friedman, Milton. 1960. "Comment" in National Bureau of Economic Research. *Demographic and Economic Change in Developed Countries*. Princeton: Princeton University Press.

Friedmann, John. 1956. "The Concept of a Planning Region", *Land Economics* 32:1-13.

Friedmann, John. 1968. "The Strategy of Deliberate Urbanization", *American Institute of Planners Journal* 34:364-373.

Greenwood, M. J. 1973. "Urban Economic Growth and Migration: Their Interaction", *Environment and Planning* 5:91-112.

Grossner, Charles. 1970. "Specialization in Canadian Cities", unpublished M.A. thesis, Department of Economics, Queen's University, Kingston.

Hanna, Frank A. 1957. "Analysis of Interstate Income Differentials: Theory and Practice" in National Bureau of Economic Research, *Regional Income*. Princeton: Princeton University Press.

Hansen, Niles M. 1967. "Development Pole Theory in a Regional Context", *Kyklos* 20:709-727.

Hansen, Niles M. 1971. *Intermediate-Size Cities as Growth Centers*. New York: Praeger Publishers.

Hansen, Niles M. 1972. "Growth Center Policy in the United States" in Niles M. Hansen (ed.), *Growth Centers in Regional Economic Development*. New York: The Free Press.

Harris, Chauncey, and Edward L. Ullman. 1945. "The Nature of Cities", *Annals of the American Academy of Political and Social Science* 242:7-17.

Harris, Seymour E. 1952. *The Economy of New England: Case Study of an Older Area*. Cambridge: The Harvard University Press.

Hartwick, John M., and Ronald W. Crowley. 1972. *Urban Economic Growth: The Canadian Case*. Working Paper A.73.5, Ministry of State for Urban Affairs, Ottawa.

Hauser, Philip M. 1965. "Urbanization: An Overview" in Philip M. Hauser and Leo F. Schnore (eds.), *The Study of Urbanization*. New York: John Wiley & Sons, Inc.

Hawley, Amos H. 1950. *Human Ecology*. New York: The Ronald Press Company.

Hawley, Amos H. 1972. "Population Density and the City", *Demography* 9:521-529.

Henripin, Jacques. 1972. *Trends and Factors of Fertility in Canada*. Statistics Canada. Ottawa: Information Canada.

Hermansen, Tormod. 1972. "Development Poles and Development Centres in National and Regional Development" in Antoni Kuklinski (ed.), *Growth Pole and Growth Centres in Regional Planning*. The Hague: Mouton and Co.

Higgins, Benjamin. 1972. "Growth Pole Policy in Canada" in Niles M. Hansen (ed.), *Growth Centers in Regional Economic Development*. New York: The Free Press.

Hilhurst, J. G. M. 1972. "Development Axes and the Diffusion of Innovation", *Development and Change* 4:1-16.

Hirschman, Albert D. 1958. The Strategy of Economic Development. New Haven: Yale University Press.

Hodge, Gerald. 1965. "The Prediction of Trade Center Viability in the Great Plains", Papers of the Regional Science Association 15:87-115.

Hodge, Gerald. 1966. The Identification of 'Growth Poles' in Eastern Ontario. Ontario Department of Economics and Development, Toronto.

Hoffmann, Walther G. 1955. British Industry, 1700-1950. Translated by W. O. Henderson and W. H. Chaloner. Oxford: Basil Blackwell.

Hoffmann, Walther G. 1958. The Growth of Industrial Economies. Translated by W. O. Henderson and W. H. Chaloner from Stadien und Typen der Industrialisierung, 1931. Manchester: Manchester University Press.

Hoover, Edgar M. 1948. The Location of Economic Activity. New York: McGraw-Hill Book Company, Inc.

Hoselitz, Bert F. 1961. "Some Problems in the Quantitative Study of Industrialization", Economic Development and Cultural Change 9:537-549.

Hoyt, Homer. 1939. The Structure and Growth of Residential Neighborhoods in American Cities. Washington, D.C.: Federal Housing Administration.

Hudson, John C. 1972. "Some Observations on Migration Theory for an Urban System". Paper presented to the International Geographical Union's Commission on Population Geography, University of Alberta, Edmonton.

Hurst, Michael E. Elliot. 1972. A Geography of Economic Behavior. Belmont, California: Wadsworth Publishing Company, Inc.

Isard, Walter. 1956. Location and Space Economy. Cambridge: The M. I. T. Press.

Isard, Walter. 1960. Methods of Regional Analysis. Cambridge: The M. I. T. Press.

Janson, Carl-Gunnar. 1969. "Some Problems of Ecological Factor Analysis" in Mattei Dogan and Stein Rokkan (eds.), Quantitative Ecological Analysis in the Social Sciences. Cambridge: The M. I. T. Press.

Keeble, D. E. 1967. "Models of Economic Development" in Richard J. Chorley and Peter Hagget (eds.), Models in Geography. London: Methuen and Co. Ltd.

King, Leslie J. 1966. "Cross-Sectional Analysis of Canadian Urban Dimensions: 1951 and 1961", Canadian Geographer 10:205-224.

Kuklinski, Antoni (ed.). 1972. Growth Pole and Growth Centres in Regional Planning. The Hague: Mouton and Co.

Kuznets, Simon. 1960. "Population Change and Aggregate Output" in National Bureau of Economic Research. Demographic and Economic Change in Developed Countries. Princeton: Princeton University Press.

Lampard, Eric E. 1968. "The Evolving System of Cities in the United States: Urbanization and Economic Development" in Harvey S. Perloff and Lowdon Wingo, Jr. (eds.), Issues in Urban Economics. Baltimore: The Johns Hopkins Press.

Lasuen, J. R. 1969. "On Growth Poles", Urban Studies 6:137-161.

Lithwick, N. H., and Gilles Paquet. 1968. "Urban Growth and Regional Contagion" in N. H. Lithwick and Gilles Paquet (eds.), Urban Studies: A Canadian Perspective. Toronto: Methuen.

Lord, Frederic M. 1963. "Elementary Models for Measuring Change" in Chester W. Harris (ed.), Problems in Measuring Change. Madison: The University of Wisconsin Press.

Losch, August. 1954. The Economics of Location. Translated from the second revised edition by William H. Woglam and Wolfgang F. Stolper. New Haven: Yale University Press.

Lowry, Ira S. 1966. Migration and Metropolitan Growth: Two Analytical Models. San Francisco: Chandler Publishing Company.

Mansfield, Edwin. 1957. "City Size and Income, 1949" in National Bureau of Economic Research, Regional Income. Princeton: Princeton University Press.

Marczewski, Jan. 1961. "Some Aspects of the Economic Growth of France", Economic Development and Cultural Change 9:369-386.

Marshall, John U. 1975. "City Size, Economic Diversity, and Functional Type: The Canadian Case", Economic Geography 51:37-49.

Martindale, Don. 1958. "Prefatory Remarks: The Theory of the City" in Max Weber, The City. Translated and edited by Don Martindale and Gertrud Neuwirth. New York: The Free Press.

Mattila, John M. and Wilbur R. Thompson. 1955. "The Measurement of the Economic Base of the Metropolitan Area", Land Economics 31:215-228.

Mattila, John M., and Wilbur R. Thompson. 1968. "Toward an Econometric Model of Urban Economic Development" in Harvey S. Perloff and Lowdon Wingo, Jr. (eds.), Issues in Urban Economics. Baltimore: The Johns Hopkins Press.

Maxwell, J. W. 1965. "The Functional Structure of Canadian Cities: A Classification of Cities", Geographical Bulletin 7:79-104.

McCalden, Gerald. 1973. "Alternative Approaches to Economic Base Analysis for Small Areas", Royal Australian Planning Institute Journal 11:91-96.

Mehtra, Surinder K. 1961. "A Comparative Analysis of the Industrial Structure of the Urban Labour Force of Burma and the United States", Economic Development and Cultural Change 9:164-190.

Meier, Richard L. 1962. A Communications Theory of Urban Growth. Cambridge: The M. I. T. Press.

Moore, C. W. 1971. "Growth Poles and Industrial Linkage Development Paths: A Research Methodology", unpublished Ph.D. dissertation, University of Washington.

Moore, C. W. 1972. "Industrial Linkage Development Paths in Growth Poles: A Research Methodology", Environment and Planning 4:253-271.

Moseley, M. J. 1973a. "The Impact of Growth Centres in Rural Regions - I. An Analysis of Spatial 'Patterns' in Brittany", Regional Studies 7:57-75.

Moseley, M. J. 1973b. "The Impact of Growth Centres in Rural Regions - II. An Analysis of Spatial 'Flows' in East Anglia", Regional Studies 7:77-94.

Moser, C. A., and Wolf Scott. 1961. British Towns. London: Oliver and Boyd.

Murdie, R. A. 1969. Factorial Ecology of Metropolitan Toronto, 1951-1961. Research Paper No. 116. Department of Geography, University of Chicago, Chicago.

Nichols, Vida. 1969. "Growth Poles: An Evaluation of their Propulsive Effect", Environment and Planning 1:193-208.

Nicholson, T. G. and M. H. Yeates. 1969. "The Ecological and Spatial Structure of the Socio-Economic Characteristics of Winnipeg, 1961", The Canadian Review of Sociology and Anthropology 6:172-178.

North, Douglass C. 1955. "Location Theory and Regional Economic Growth", Journal of Political Economy 63: 243-258.

North, Douglass C. 1956. "A Reply", Journal of Political Economy 64:165-168.

North, Douglass C. 1961. The Economic Growth of the United States 1790-1860. Englewood Cliffs, N. J.: Prentice-Hall, Inc.

Ostry, Sylvia. 1968. The Female Worker in Canada. 1961 Census Monograph. Catalogue Number 99-553. Dominion Bureau of Statistics. Ottawa: The Queen's Printer.

Parai, Louis. 1974. The Economic Impact of Immigration. Manpower and Immigration. Ottawa: Information Canada.

Patel, Surendra J. 1961. "Rates of Industrial Growth in the Last Century, 1860 - 1958", Economic Development and Cultural Change 9:316-330.

Perloff, Harvey S. 1957. "Interrelations of State Income and Industrial Structure", Review of Economics and Statistics 39:162-171.

Perloff, Harvey S., Edgar S. Dunn Jr., Eric E. Lampard, and Richard F. Muth. 1960. Regions, Resources and Economic Growth. Baltimore: The Johns Hopkins Press.

Perloff, Harvey, and Lowdon Wingo Jr. 1964. "Natural Resource Endowment and Regional Economic Growth" in John Friedmann and William Alonso (eds.), Regional Development and Planning. Cambridge: The M. I. T. Press.

Perroux, Francois. 1964. "Economic Space: Theory and Applications" in John Friedmann and William Alonso (eds.), Regional Development and Planning. Cambridge: The M. I. T. Press.

Peterson, William. 1975. Population. Third Edition. New York: Macmillan Publishing Co., Inc.

Pred, Allan R. 1966. The Spatial Dynamics of U.S. Urban-Industrial Growth, 1800-1914. Cambridge: The M. I. T. Press.

Putnam, D. F., and R. G. Putnam. 1970. Canada: A Regional Analysis. Toronto: J. M. Dent & Sons Limited.

Ray, D. Michael. 1968. "Urban Growth and the Concept of Functional Region" in N. H. Lithwick and Gilles Paquet (eds.), Urban Studies: A Canadian Perspective. Toronto: Methuen.

Ray, D. Michael. 1969. "The Spatial Structure of Economic and Cultural Differences: A Factorial Ecology of Canada", Papers of the Regional Science Association 23:7-23.

Ray, D. Michael. 1971. Dimensions of Canadian Regionalism. Geographic Paper Number 49. Department of Energy, Mines and Resources. Ottawa: Information Canada.

Reissman, Leonard. 1964. The Urban Process. New York: The Free Press.

Richardson, Harry W. 1969. Regional Economics. London: Weidenfield and Nicolson.

Richardson, Harry W. 1971. Urban Economics. Baltimore: Penguin Books, Inc.

Rostow, W. W. 1956. "The Take-off into Self-Sustained Growth", Economic Journal 66:25-48.

Royal Commission on Canada's Economic Prospects. Final Report. 1957. Ottawa: The Queen's Printer.

Rummel, R. J. 1970. Applied Factor Analysis. Evanston: Northwestern University Press.

Samuelson, Paul A. 1955. Economics. Third Edition. New York: McGraw-Hill Book Company, Inc.

Shryock, Henry S., Jacob Siegel, and Associates. 1973. The Methods and Materials of Demography. Second Printing (rev.). United States Bureau of the Census. United States Government Printing Office, Washington, D.C.

Siebert, Horst. 1969. Regional Economic Growth: Theory and Policy. Scranton, Pennsylvania: International Textbook Company.

Sjoberg, Gideon. 1965. "The Origin and Evolution of Cities", Scientific American 213:55-62.

Spengler, Joseph J. 1949. "Theories of Socio-Economic Growth" in Problems in the Study of Economic Growth. New York: National Bureau of Economic Research.

Stafford, James. 1975. "Census Indicators of Urban Development in Alberta". Paper presented at Conference on Social Indicators, University of Alberta, Edmonton.

Statistics Canada. 1973. 1971 Census of Canada. Catalogue Number 92-702. Ottawa: Information Canada.

Stigler, George J. 1956. Trends in Employment in the Service Industries. National Bureau of Economic Research, Number 59, General Series. Princeton: Princeton University Press.

Stone, Leroy O. 1967. Urban Development in Canada. 1961 Census Monograph. Dominion Bureau of Statistics. Ottawa: The Queen's Printer.

Stouffer, Samuel A. 1960. "Intervening Opportunities and Competing Migrants", Journal of Regional Science 2:1-26.

Sultan, Paul. 1957. Labour Economics. New York: Henry Holt and Company.

Sweet, James A. 1970. "Family Composition and the Labor Force Activity of American Wives", Demography 7:195-209.

Tattersall, James N. 1962. "Exports and Economic Growth: The Pacific Northwest: 1880-1960", Regional Science Association Papers and Proceedings 9:215-234.

Thomas, Morgan D. 1963. "Regional Economic Growth and Industrial Development", Regional Science Association Papers 10:61-76.

Thomas, Morgan D. 1964. "The Export Base and Development Stages Theories of Regional Economic Growth: An Appraisal", Land Economics 40:421-432.

Thomlinson, Ralph. 1969. Urban Structure. New York: Random House.

Thompson, Wilbur R. 1965. "Urban Economic Growth and Development in a National System of Cities" in P. M. Hauser and L. F. Schnore (eds.), The Study of Urbanization. New York: John Wiley and Sons.

Thompson, Wilbur R. 1968. "Internal and External Factors in the Development of Urban Economics" in Harvey S. Perloff and Lowdon Wingo, Jr. (eds.), Issues in Urban Economics. Baltimore: The Johns Hopkins Press.

Tiebout, Charles M. 1956. "Exports and Regional Economic Growth", Journal of Political Economy 64:160-164.

Tiebout, Charles M. 1962. The Community Economic Base Study. Supplementary Paper Number 16. Published by the Committee for Economic Development, New York.

Tisdale, Hope. 1942. "The Process of Urbanization", Social Forces 20:311-316.

Ullman, Edward L., and Michael F. Dacey. 1960. "The Minimum Requirements Approach to the Urban Economic Base", Papers and Proceedings of the Regional Science Association 6:175-194.

Ullman, Edward L., Michael F. Dacey, and Harold Brodsky. 1971. The Economic Base of American Cities. Seattle: The University of Washington Press.

Vanneste, Olivier. 1971. The Growth Pole Concept and the Regional Economic Policy. Brugge: De Tempel.

Warkentin, John. 1968. Canada: A Geographical Interpretation. Toronto: Methuen.

Weber, Adna Ferrin. 1963. The Growth of Cities in the Nineteenth Century. Ithaca: Cornell University Press.

Weber, Alfred. 1929. Theory of the Location of Industries. Translated by Carl J. Freidrich from Ueber den Studort der Industrien, 1909. Chicago: University of Chicago Press.

Weber, Max. 1958. The City. Translated and edited by Don Martindale and Gertrud Neuwirth. New York: The Free Press.

Wingo, Lowdon. 1973. "The Quality of Life: Toward a Micro-economic Definition", Urban Studies 10:3-18.

Wirth, Louis. 1938. "Urbanization as a Way of Life", American Journal of Sociology 44:1-24.

Wonnacott, Ronald J., and Thomas H. Wonnacott. 1970. Econometrics. New York: John Wiley & Sons, Inc.

Woytinsky, W. S., and E. S. Woytinsky. 1953. World Population and Production. New York: The Twentieth Century Fund.

Yeates, Maurice H., and Barry J. Gardner. 1971. The North American City. New York: Harper and Row.

Zipf, G. K. 1949. Human Behaviour and the Principle of Least Effort. Reading, Mass.: Addison-Wesley.

Appendices

Appendix A

Industry Classification

- Agriculture
- Forestry and Logging
- Fishing and Trapping
- Mining
- Manufacturing
  - Food Products
  - Liquors and Beverages
  - Tobacco Products
  - Rubber Products
  - Leather Goods
  - Textile Products
  - Hosiery and Knit Goods
  - Clothing
  - Lumber and Wood Products
  - Furniture
  - Pulp and Paper
  - Printing and Publishing
  - Metals and Machinery
    - Primary Iron and Steel
    - Non-ferrous Metal Smelting and Refining
    - Metal Fabricating

    - Agricultural Implements
    - Other Machinery
  - Transportation Equipment
    - Ship and Boat Building
    - Railway Equipment
    - Automobiles, Bicycles and Carriages
    - Aircraft
  - Electrical Apparatus
  - Non-metallic Mineral Products
  - Petroleum Refining
  - Chemicals
  - Miscellaneous Manufacturing
- Electricity and Gas
- Construction
- Transportation
  - Air Transportation
  - Railway, Express and Telegraph
  - Water Transportation
  - Other Transportation
  - Communication
  - Storage
- Trade
  - Wholesale Trade
  - Retail Trade

- Finance, Insurance, and Real Estate
  - Finance
  - Insurance and Real Estate
- Services
  - Education and Related Services
  - Health and Welfare Services
  - Religious Services
  - Food and Lodging
  - Recreation
  - Personal Services
  - Business Services
  - Repair Services
  - Other Services
- Public Administration
  - Federal Administration
    - Defence Services
    - Post Office
    - Other Federal Administration
  - Provincial Administration
  - Local Administration

SOURCE: Grossner, 1970, pp. 100-101

Appendix B

Description of Urban Areas

Urban Unit	Geographic Area
Newfoundland	
St. John's	Urban Area
Nova Scotia	
Halifax	Halifax, Dartmouth
Sydney	Sydney, Glace Bay, New Waterford
New Brunswick	
Moncton	U. A.
Saint John	U. A.
Quebec	
Chicoutimi	Chicoutimi, Jonquiere, Arvida
Drummondville	U. A.
Granby	U. A.
Montreal	Metropolitan Area
Quebec	M. A.
St. Jean	U. A.
Shawinigan Falls	Shawinigan Falls, Grand Mère
Sherbrooke	U. A.
Trois Rivières	Trois Rivières, Cap de la Madeleine
Valleyfield	U. A.

Ontario		
	Belleville	U. A.
	Brantford	U. A.
	Chatham	U. A.
	Cornwall	U. A.
	Guelph	U. A.
	Hamilton	M. A.
	Kingston	U. A.
	Kitchener	Galt, Kitchener, Waterloo
	London	U. A.
	Oshawa	U. A.
	Ottawa	M. A.
	Peterborough	U. A.
	Port Arthur	Fort William, Port Arthur
	St. Catherines	U. A.
	Sarnia	U. A.
	Sault Ste. Marie	U. A.
	Sudbury	U. A.
	Timmins	U. A.
	Toronto	M. A.
	Welland	U. A.
	Windsor	M. A.
Manitoba		
	Brandon	U. A.
	Winnipeg	M. A.

Saskatchewan	
Moose Jaw	U. A.
Regina	U. A.
Saskatoon	U. A.
Alberta	
Calgary	M. A.
Edmonton	M. A.
Lethbridge	U. A.
British Columbia	
Vancouver	M. A.
Victoria	U. A.

Appendix C

Independent Variables and Data Sources

	Variable*					
	3	4	6	7	8	9
Belleville	1.91	39.3	-0.14	2.63	-3.12	-12.43
Brandon	2.20	22.1	-1.28	0.68	-1.55	-0.43
Brantford	12.83	52.8	-4.59	-1.34	1.04	1.08
Calgary	2.52	32.6	2.34	-1.57	-1.97	-3.78
Chatham	5.95	54.3	-3.01	0.23	0.12	5.06
Chicoutimi	41.75	75.2	1.22	-0.13	0.32	7.14
Cornwall	8.64	5.0	-2.66	-2.34	-0.76	7.39
Drummondville	4.06	5.0	9.17	3.42	-3.73	-1.41
Edmonton	2.10	26.1	-1.16	2.07	-3.33	5.11
Granby	6.85	4.5	-3.83	5.39	-3.11	2.80
Guelph	2.33	35.6	-2.00	-0.48	1.08	-4.92
Halifax	7.97	29.0	-5.49	0.47	2.37	4.45
Hamilton	2.59	26.2	2.44	-1.17	-0.82	-1.97
Kingston	2.41	51.5	-2.34	0.00	1.24	3.94
Kitchener	5.86	19.7	-2.66	0.50	-0.42	2.02
Lethbridge	2.13	14.6	-1.50	1.79	2.09	-1.13
London	2.59	34.5	-1.75	-0.46	1.60	-7.33
Moncton	3.11	43.6	-0.25	-1.52	1.27	2.97
Montreal	2.69	30.2	-0.62	1.99	-1.80	1.38
Moose Jaw	3.05	12.7	-2.16	-0.48	3.12	-1.08
Oshawa	3.40	82.3	-0.00	0.77	0.18	9.66
Ottawa	4.87	2.4	-1.68	-1.78	1.29	-3.79
Peterborough	3.42	15.2	3.15	-3.53	3.96	7.09
Port Arthur	9.77	36.7	-4.81	-1.71	-2.28	-10.22
Quebec	3.60	17.7	1.00	-0.00	1.14	4.95
Regina	2.73	16.3	-0.77	1.03	-4.13	6.19
St. Catherines	6.02	61.5	0.51	1.40	0.00	7.20
St. Jean	1.88	47.9	-0.99	1.36	-0.53	-28.93
St. John's	3.01	20.4	-1.41	-3.96	3.86	-0.00
Saint John	4.02	29.1	-0.95	3.03	-2.58	8.07
Sarnia	7.17	59.2	7.29	-0.63	1.12	7.17
Saskatoon	1.62	20.3	-1.22	-0.36	-1.15	-8.07
Sault Ste. Marie	5.60	82.1	15.51	-0.87	-1.54	7.93
Shawinigan	3.42	45.5	-2.28	3.90	-0.58	0.22
Sherbrooke	5.53	22.7	-1.02	-1.72	3.57	-0.12
Sudbury	12.45	70.3	3.57	0.24	0.24	8.43
Sydney	5.55	82.6	5.11	-0.30	2.25	5.42
Timmins	15.59	11.2	-3.58	-0.22	-2.34	-2.86
Toronto	2.11	31.8	-1.16	1.06	-0.00	3.17
Trois Rivières	4.48	19.9	0.00	4.38	-2.38	-2.44
Valleyfield	4.07	19.4	-1.47	-8.11	3.27	-0.93
Vancouver	6.50	22.8	-0.49	-0.34	0.40	0.00
Victoria	2.82	26.6	-1.40	-0.33	0.44	2.31
Welland	2.69	62.3	-1.41	2.66	-3.32	13.04
Windsor	5.25	80.4	8.81	-5.49	5.14	-39.79
Winnipeg	2.34	35.7	0.00	-0.16	0.30	-2.60

*Variables listed at end of this appendix

Appendix C (continued)

	Variable*				
	10	11	12	14	15
Belleville	0.37	-0.37	-0.45	306	73.8
Brandon	1.33	-0.49	-0.37	131	-135.5
Brantford	-0.70	0.64	-0.37	437	-100.9
Calgary	0.91	3.41	0.34	137	-8.7
Chatham	-0.11	-0.08	-0.38	325	-0.6
Chicoutimi	-0.32	0.27	-0.26		
Cornwall	-0.05	0.26	-0.41	855	-122.7
Drummondville	-0.20	0.62	-0.44	848	-281.5
Edmonton	1.50	3.47	0.35	134	5.4
Granby	-0.40	-0.23	-0.75	298	167.0
Guelph	-0.53	0.21	-0.49	251	97.0
Halifax	1.69	0.26	-0.08	150	76.1
Hamilton	-1.07	-1.04	0.35	548	192.8
Kingston	0.39	-0.28	-0.34	295	-8.0
Kitchener	-0.84	0.07	-0.39	383	48.2
Lethbridge	0.92	0.01	-0.45	129	-11.5
London	0.26	1.17	-0.05	315	103.6
Moncton	1.26	0.13	-0.16	253	-125.1
Montreal	0.26	-1.79	4.34	236	22.3
Moose Jaw	0.99	-0.17	-0.42	180	1.4
Oshawa	-1.77	0.25	-0.42		
Ottawa	1.27	0.68	0.33	199	46.0
Peterborough	-0.88	0.02	-0.38	441	190.0
Port Arthur	0.36	-0.69	-0.27	405	-54.3
Quebec	0.44	-2.44	0.42	156	24.9
Regina	1.14	-1.07	-0.26	135	87.4
St. Catherines	-1.01	0.90	-0.32	594	-74.2
St. Jean	-0.23	0.24	-0.52	340	-3.9
St. John's	1.23	-0.56	-0.15	112	3.9
Saint John	1.28	-1.09	-0.25	148	228.1
Sarnia	-1.31	0.15	-0.37	1431	941.3
Saskatoon	1.13	-0.06	-0.26	146	-31.8
Sault Ste. Marie	-1.43	0.46	-0.16		
Shawinigan	-0.99	-0.54	-0.61	974	7.8
Sherbrooke	0.00	-0.32	-0.34	361	-40.3
Sudbury	-1.41	0.70	-0.44		
Sydney	-1.07	-0.10	-0.28		
Timmins	-0.74	-0.66	-0.70	85	-85.6
Toronto	0.03	3.40	3.94	231	50.4
Trois Rivières	-0.12	-0.74	-0.47	541	70.9
Valleyfield	-1.27	0.32	-0.44	576	-138.8
Vancouver	0.66	-1.18	1.98	187	-55.7
Victoria	1.04	-1.71	-0.07	105	-39.0
Welland	-1.20	0.38	[redacted]	1223	-673.4
Windsor	-1.59	-1.67	0.31	[redacted]	[redacted]
Winnipeg	0.78	-0.75	0.67	169	-18.3

*Variables listed at end of this appendix

Appendix C (continued)

3. 1951 location quotient coefficient: Grossner, 1970, Table 12.

4. Percentage of 1951 manufacturing labour force engaged in capital-goods industries: computed from Dominion Bureau of Statistics, 1953, Tables 17, 21.

6. Residual change in location quotient coefficient: computed from Grossner, 1970, Table 12.

7. Residual change in percentage of the labour force engaged in manufacturing industries: computed from Dominion Bureau of Statistics, 1953, Tables 17, 21; Dominion Bureau of Statistics, 1963c, Table 2; Dominion Bureau of Statistics, 1963d, Table 5; Dominion Bureau of Statistics, 1964a, Table 6.

8. Residual change in percentage of the labour force engaged in service industries: computed from Dominion Bureau of Statistics, 1953, Tables 17, 21; Dominion Bureau of Statistics, 1963c, Table 2; Dominion Bureau of Statistics, 1963d, Table 5; Dominion Bureau of Statistics, 1964a, Table 6.

9. Residual change in percentage of the manufacturing labour force engaged in capital-goods industries: computed from Dominion Bureau of Statistics, 1953, Tables 17, 21; Dominion Bureau of Statistics, 1963c, Table 2; Dominion Bureau of Statistics, 1963d, Table 5; Dominion Bureau of Statistics, 1964a, Table 6.

10. Factor score summarizing percentages of the 1951 labour force engaged in manufacturing and service industries: computed from Dominion Bureau of Statistics, 1953, Tables 17, 21.

11. Factor score summarizing residual change in labour force size and in number of retail, wholesale, and service units: computed from Dominion Bureau of Statistics, 1953, Tables 8, 14; Dominion Bureau of Statistics, 1954a, Tables 2, 3; Dominion Bureau of Statistics, 1954b, Tables 4, 5, 22, 23; Dominion Bureau of Statistics, 1963a, Table 7a; Dominion Bureau of Statistics, 1964b, Tables 2, 6; Dominion Bureau of Statistics, 1965b, Tables 33, 34; Dominion Bureau of Statistics, 1966c, Tables 6, 7.

12. Factor score summarizing labour force size and number of retail, wholesale, and service units in 1951: computed from Dominion Bureau of Statistics, 1953, Tables 8, 14; Dominion Bureau of Statistics, 1954a, Tables 2, 3; Dominion Bureau of Statistics, 1954b, Tables 4, 5, 22, 23.

14. 1951 average value added (in $000's): computed from Dominion Bureau of Statistics, 1954c, Tables 17, 18.

15. Residual change in average value added (in $000's): computed from Dominion Bureau of Statistics, 1954c, Tables 17, 18; Dominion Bureau of Statistics, 1965a, Tables 65, 74.